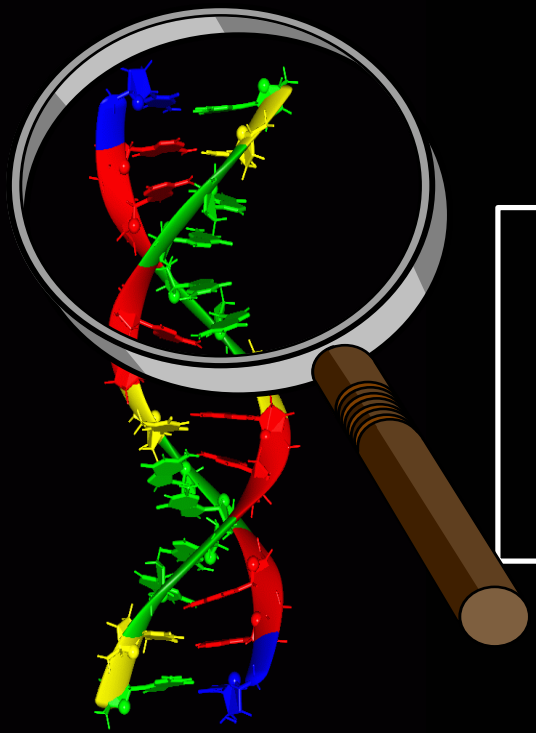
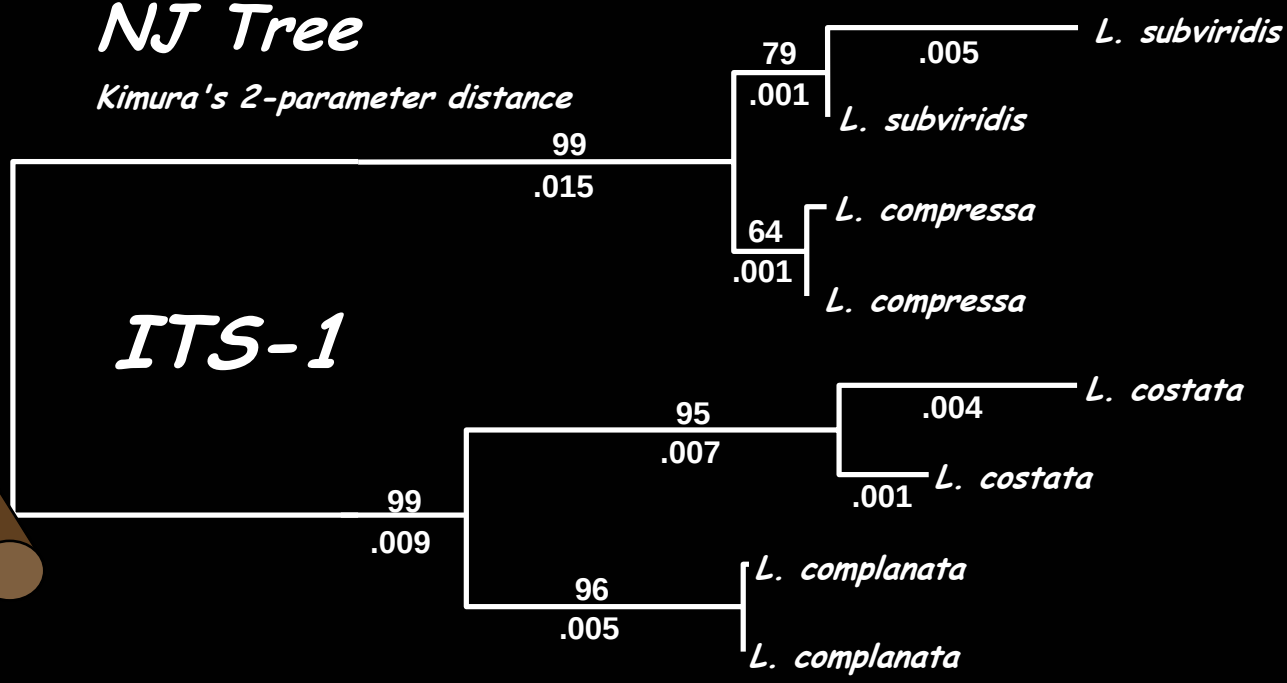
Phylogenetics and Molecular Systematics

Tim L. King
and
John Switzer

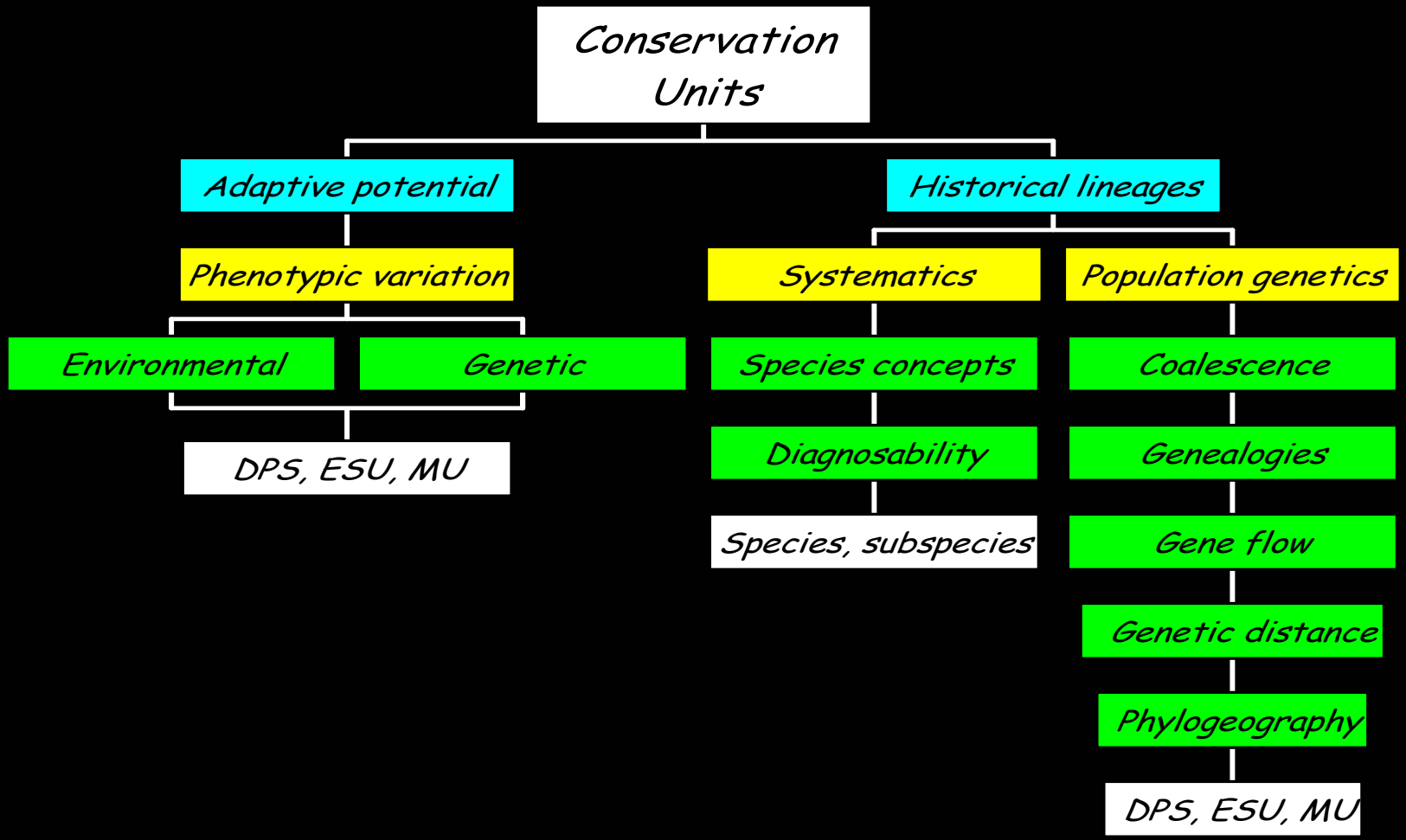
Leetown Science Center





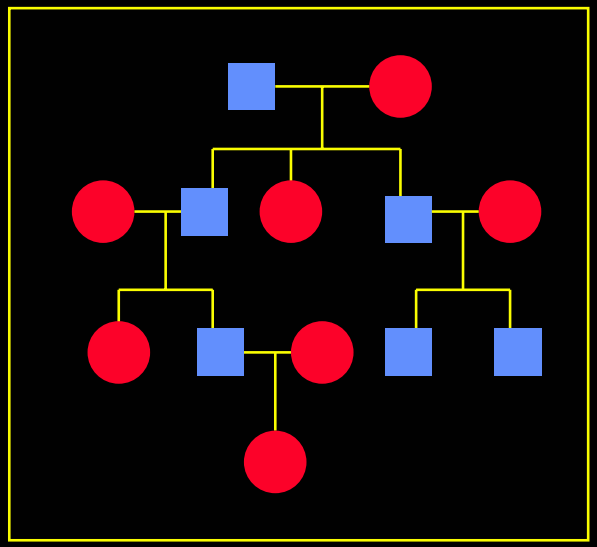
USGS
science for a changing world

Delineation of Units for Conservation and Management

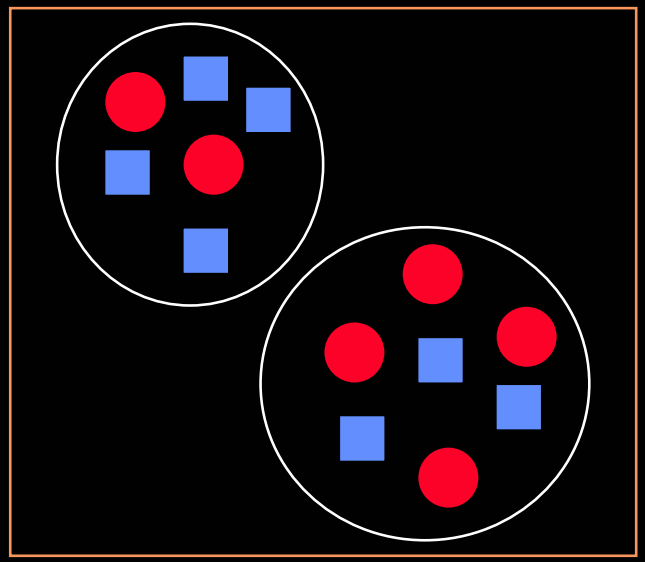


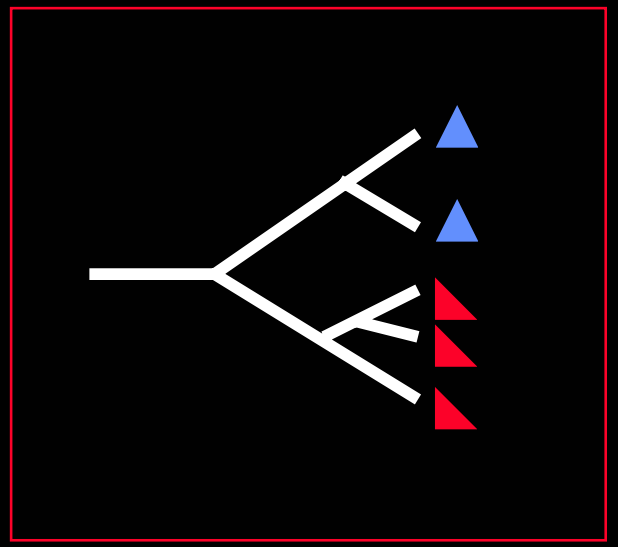
Individual-Population-Species Continuum

"The stream of heredity makes phylogeny" *G. G. Simpson, 1945*



"Family tree"





"Phylogenetic tree"

Fine scale (microevolutionary)

Broad scale (macroevolutionary)

USGS science for a changing world

Individual-Population-Species Continuum

- DNA sequence analysis: targets taxonomic or phylogeographic (i.e., Species) end of the spectrum.
  [Broad-brush]

- Microsatellite DNA and allozyme analyses: targets the Individual-Population end.
  [Fine-brush]

USGS
science for a changing world

Molecular Systematics

Detection, description, and explanation of molecular diversity, both within and among species

systematics + evolutionary theory + molecular genetics

Reconstructing evolutionary history based on shared attributes of extant and fossil organisms

Molecular Systematic Questions

- Geographic Origin(s)
- Species Relatedness
- Species Status
- Classification
- Conservation Priorities
- Evolutionary Novelty

USGS
science for a changing world

Systematic Techniques: Molecular

- Nucleotide sequence data - genealogical data!
- Restriction site data - genealogical data
- SNPs - genealogical data (haplotype determination)
- Restriction fragment data - not
- Microsatellite data - not
- AFLP data - not
- RAPDs data – not

Gene genealogy -

USGS

Overview

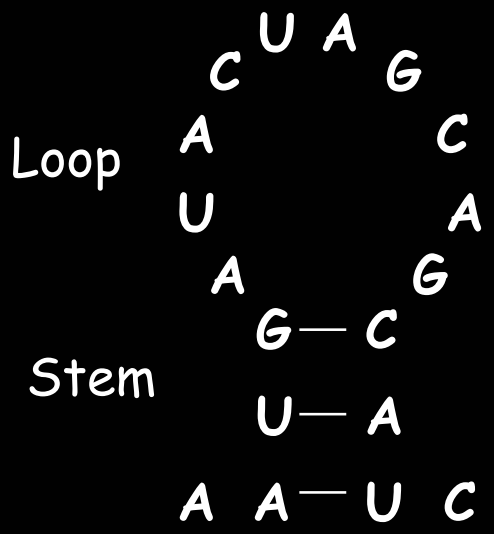
- Molecular Tools: DNA sequences
  - Changes at the DNA level
  - Choosing the best gene(s)
  - Molecular clocks
- Broad brush: Molecular Systematics
  - Interpretation
- Medium Brush: Phylogeography
  - Interpretation
  - Example: crystal darter

USGS
science for a changing world

Changes at the DNA level

Types of Genes

Structural (rDNA)	Protein coding	Non-coding



RNA secondary structure

DNA: C A A C G T

Transcription ↓ ↓

mRNA: G U U G C A

Translation ↓ ↓

Protein: Valine Alanine

Introns

Repetitive DNA

?

Molecular Tools: DNA sequences

USGS science for a changing world

Changes at the DNA level

Types of Nucleotide Substitutions

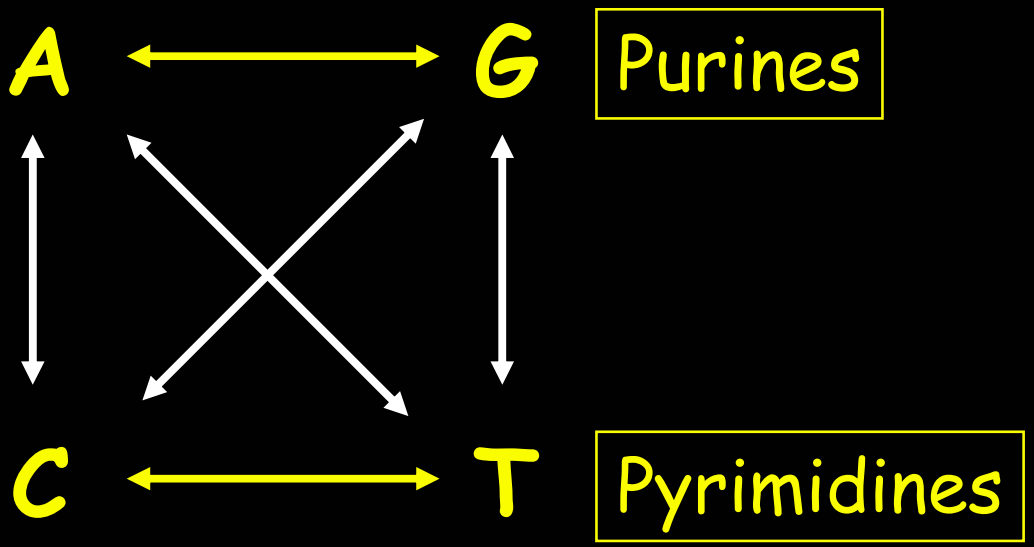
1	2	3	1	2	3
C	A	A	C	G	T
		C			C
	↓	T		↓	A
		G			G
Valine			Alanine		

- Synonymous (silent)
  - No change in amino acid
  - Most 3rd position changes
- Nonsynonymous (replacement)
  - Changes amino acid coded for
  - Most first and all second position changes

USGS science for a changing world

Changes at the DNA level

Types of Nucleotide Substitutions



- Transitions
  - Purine ↔ Purine
  - Pyrimidine ↔ Pyrimidine
- Transversions
  - Purine ↔ Pyrimidine
  - Pyrimidine ↔ Purine

USGS science for a changing world

Changes at the DNA level

Modeling DNA Substitutions

- Huge amount of sequence data available
- Realistic models of substitutions at different classes of sites
- Models can be incorporated into any of the phylogeny-building methods
- Leads to stronger interpretations of data

USGS
science for a changing world

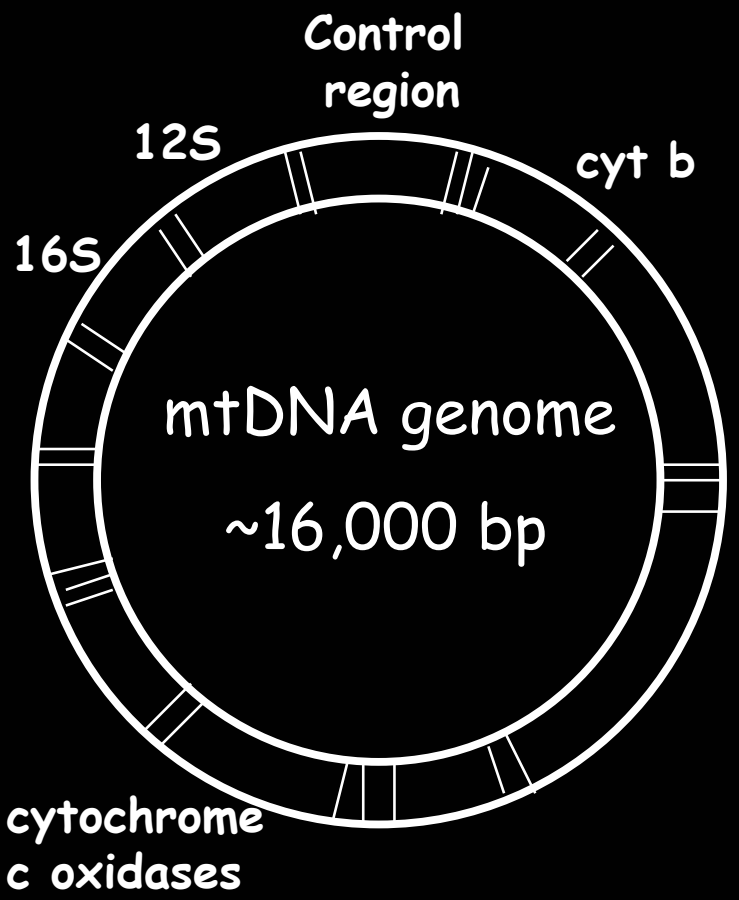
Choosing the best gene(s)

	mtDNA	Nuclear DNA
Inheritance	Maternal	Bi-parental
Ploidy	Haploid	Diploid
Ne	1/4 X	X
Sensitivity		
Broad	*(genome rearragements)	***(rDNA-stems)
Medium	**(rDNA)	*(rDNA-loops)
Fine	***(proteins, D-loop)	**(introns)

USGS science for a changing world

Choosing the best gene(s)

mitochondrial DNA



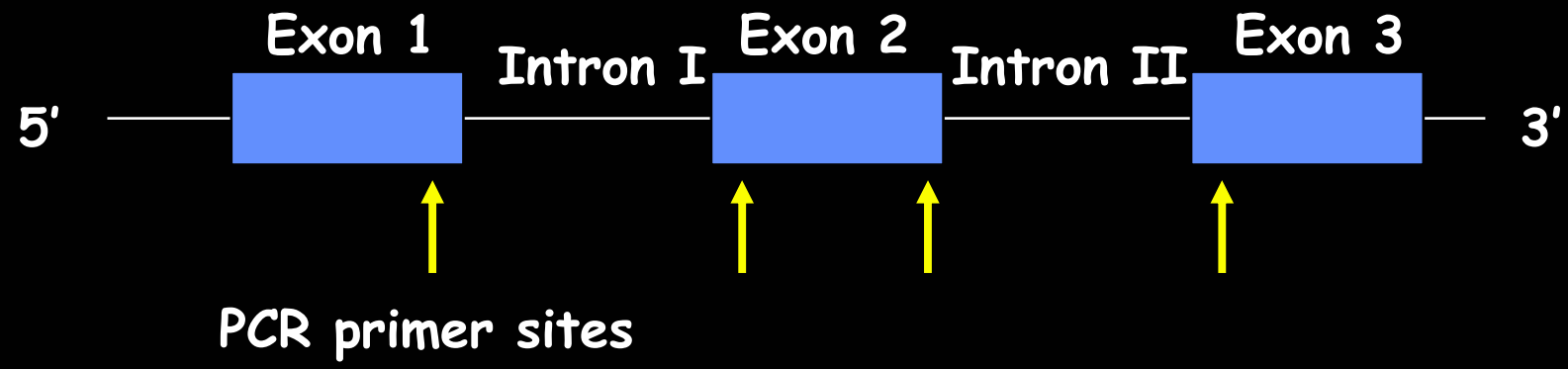
- Small, compact genome
  - 2 rDNAs
  - 13 protein coding genes
  - 22 tRNAs
  - Gene order "stable"- mostly
  - Most individuals- one sequence
- "Universal primers" (Kocher et al, 1989)
  - Extensive intraspecific polymorphism
  - Sequences available for many taxa
  - Connectivity between studies
  - RFLPs of amplified genes common

USGS
science for a changing world

Choosing the best gene(s)

nuclear introns

- Schematic of a typical nuclear protein coding gene



- Primers = exon primed, intron crossing, or EPIC
- Intron sizes can be highly variable among species
  - difficult to predict PCR product size
- Pseudogenes- amplify part of exon also to confirm

USGS
science for a changing world

Molecular Clocks
Generalizations

- Can be used to date divergences, but remember:
- Usually rely on biogeographic or fossil data
- Not like timepiece- "stochastically constant"
  - Most 'accurate' averaged over long time periods
- Can have different rates for:
  - Different lineages
  - Different genes
  - Different categories of sites along a gene
- Better than nothing!?!

USGS science for a changing world

Molecular Systematics (Phylogenetics)

- Attempt to reconstruct *historical relationships* among organisms
  - Use of state-of-the-art methods to look into the past

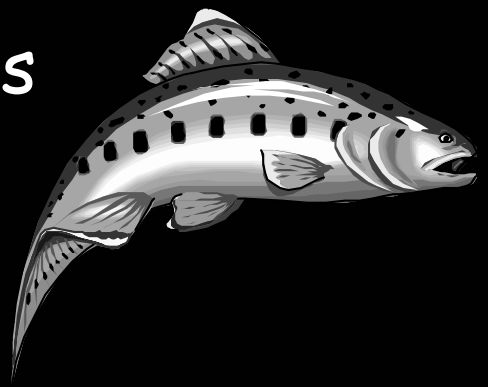
- Draws from:
  - Molecular genetics (tools)
  - Molecular evolution (how tools work)
  - Statistical hypothesis testing (how to address questions)

- *Most common uses* –
  - Resolving taxonomic uncertainties
  - Defining management units

USGS
science for a changing world

Traditional vs. Molecular Systematics

- Traditional
  - Direct observation of shared attributes
    - morphology (fossils too)
    - physiology
    - development

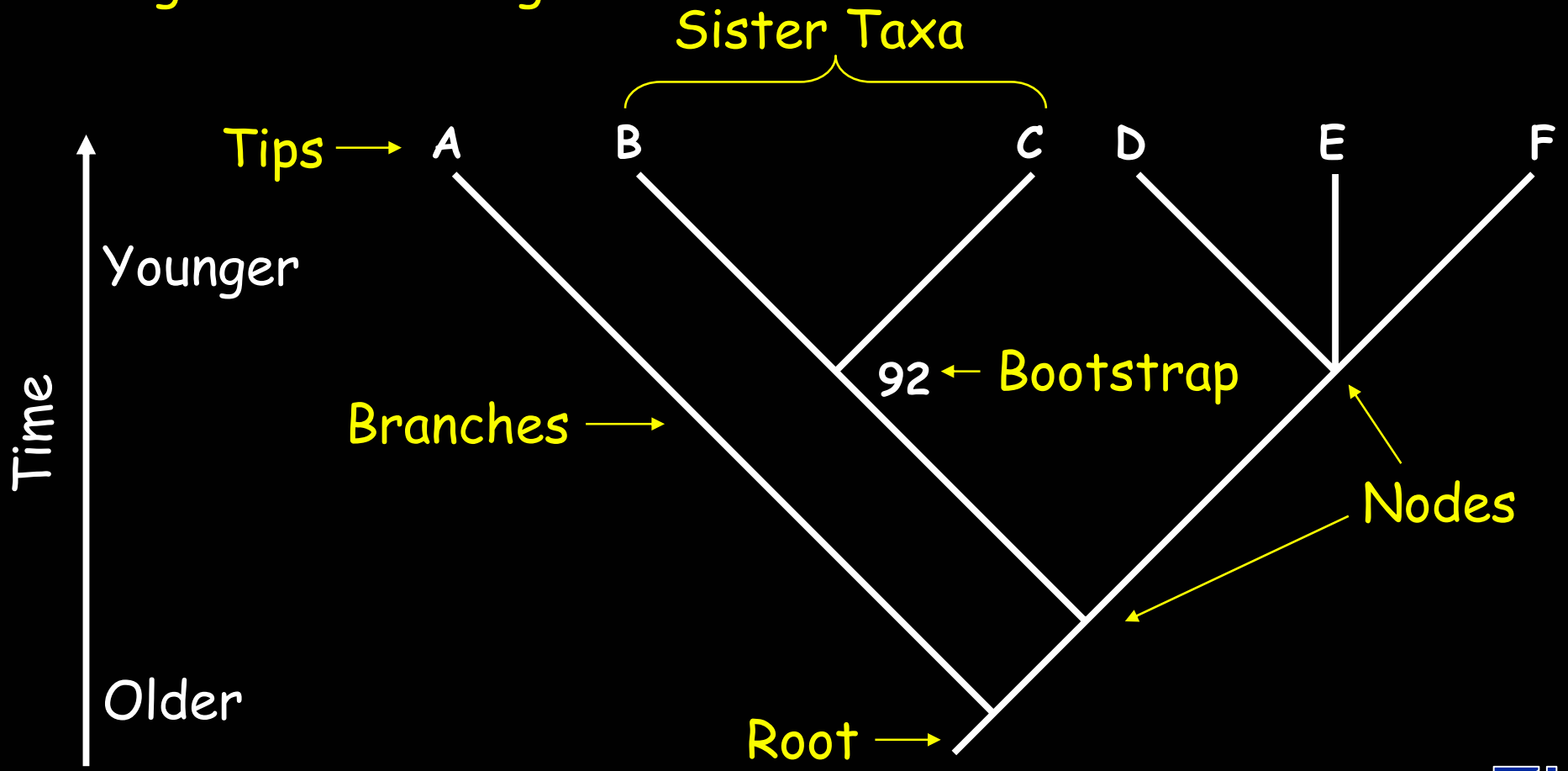


- Molecules
  - Shared attributes = mostly DNA sequence data
    - Approach can be used even when morphological differences few or misleading
    - Number of characters to compare virtually unlimited!

USGS
science for a changing world

"Tree Thinking"

Cladogram: branching order

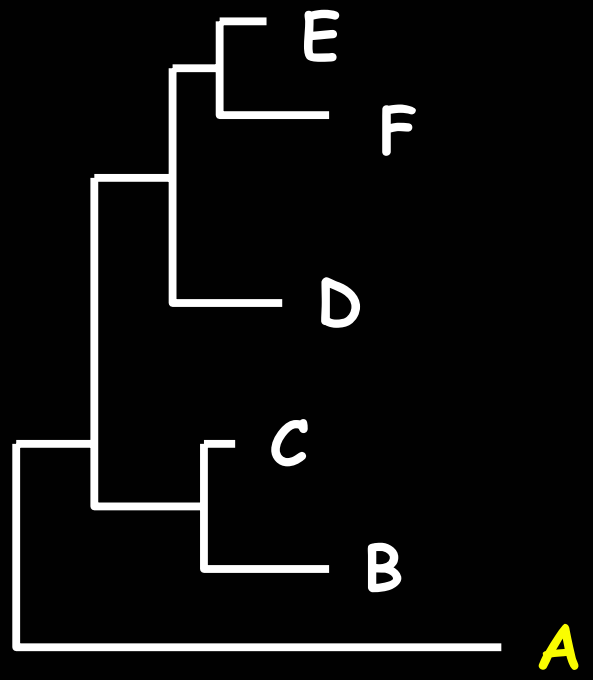


USGS
science for a changing world

"Tree Thinking"

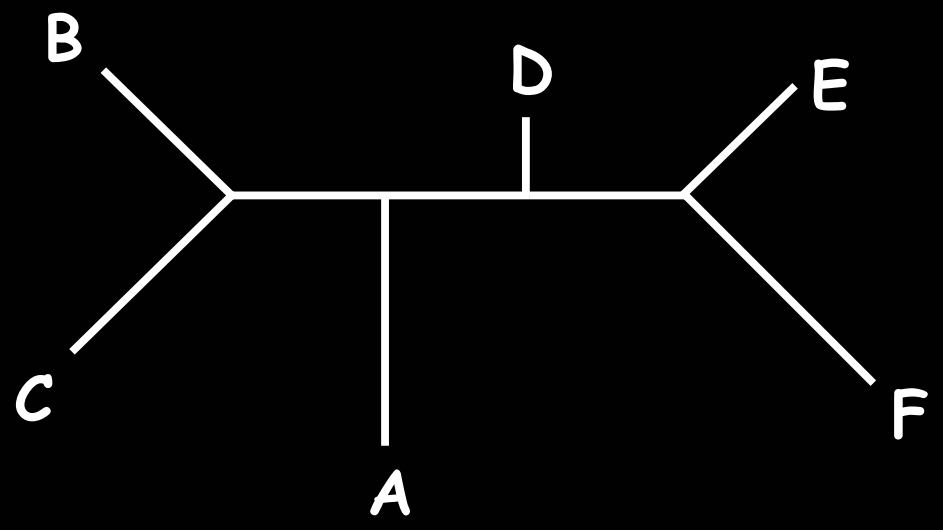
Phylogram: branching order and scaled branch lengths

Rooted Phylogram



— 0.5% sequence divergence

Unrooted Phylogram

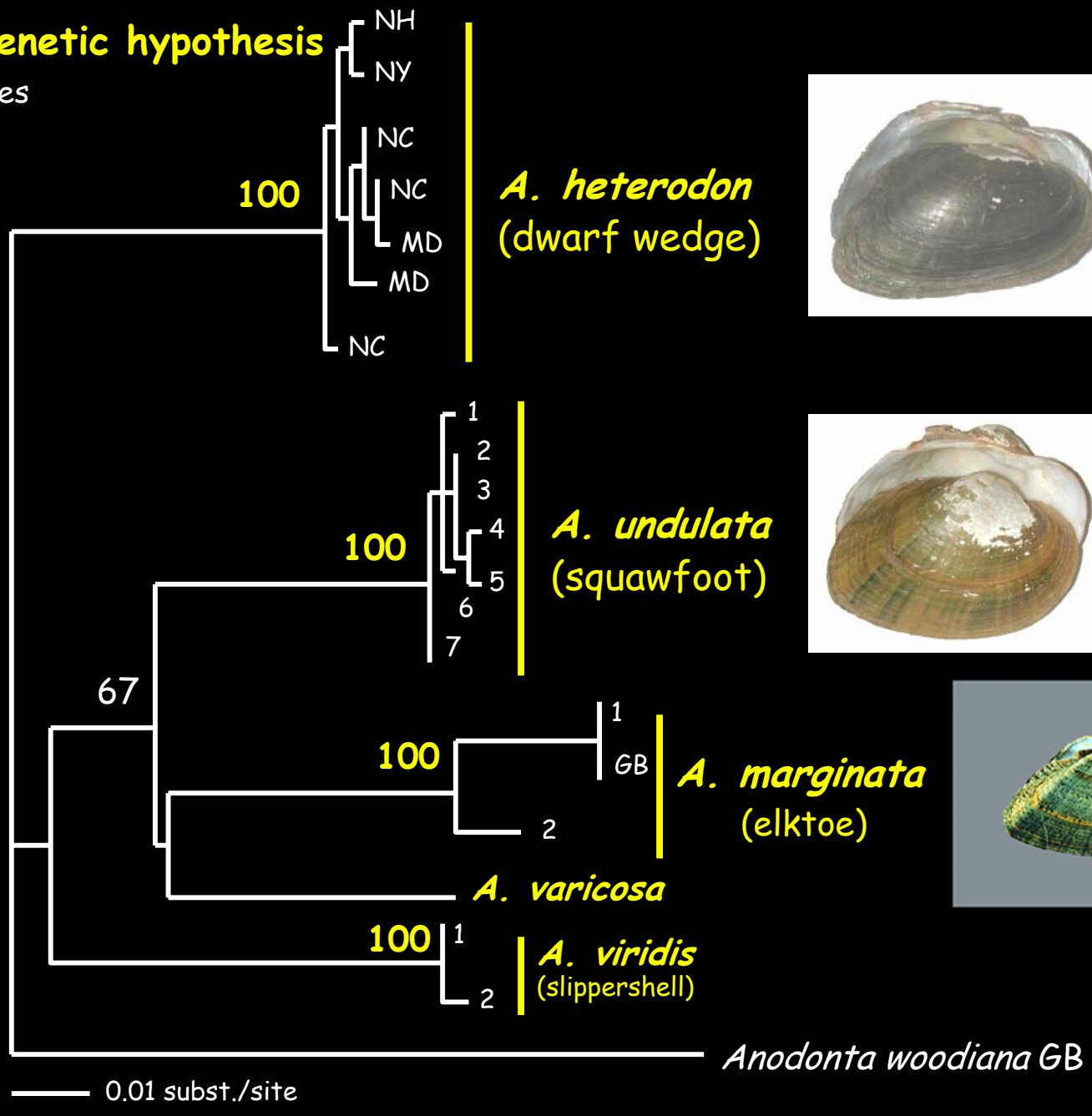


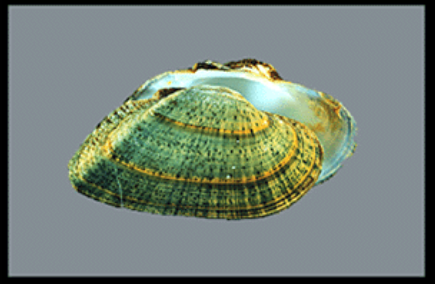
— Genetic distance

USGS
science for a changing world

Alasmidonta phylogenetic hypothesis

COI: NJ K2P distances

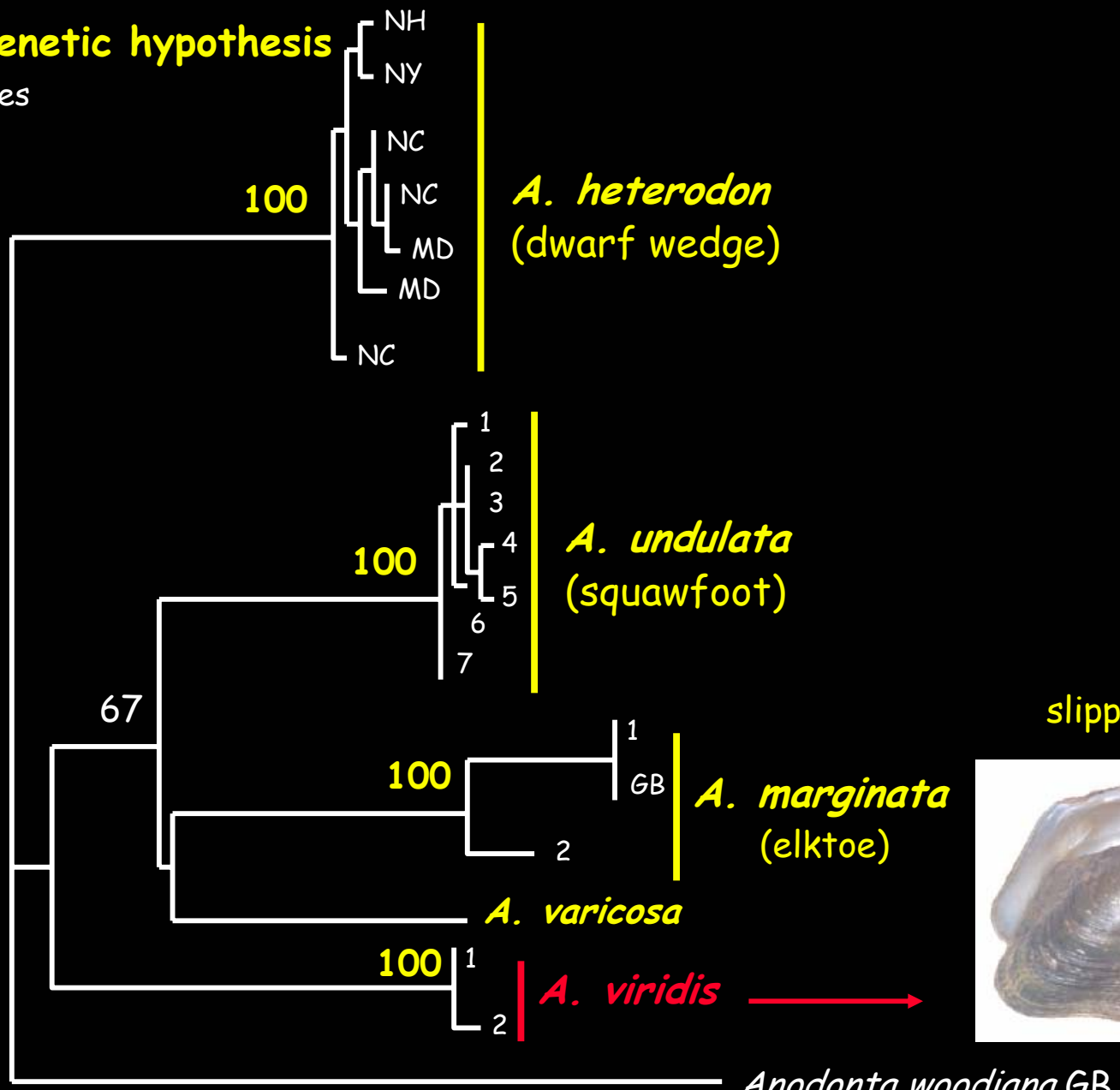




USGS
science for a changing world

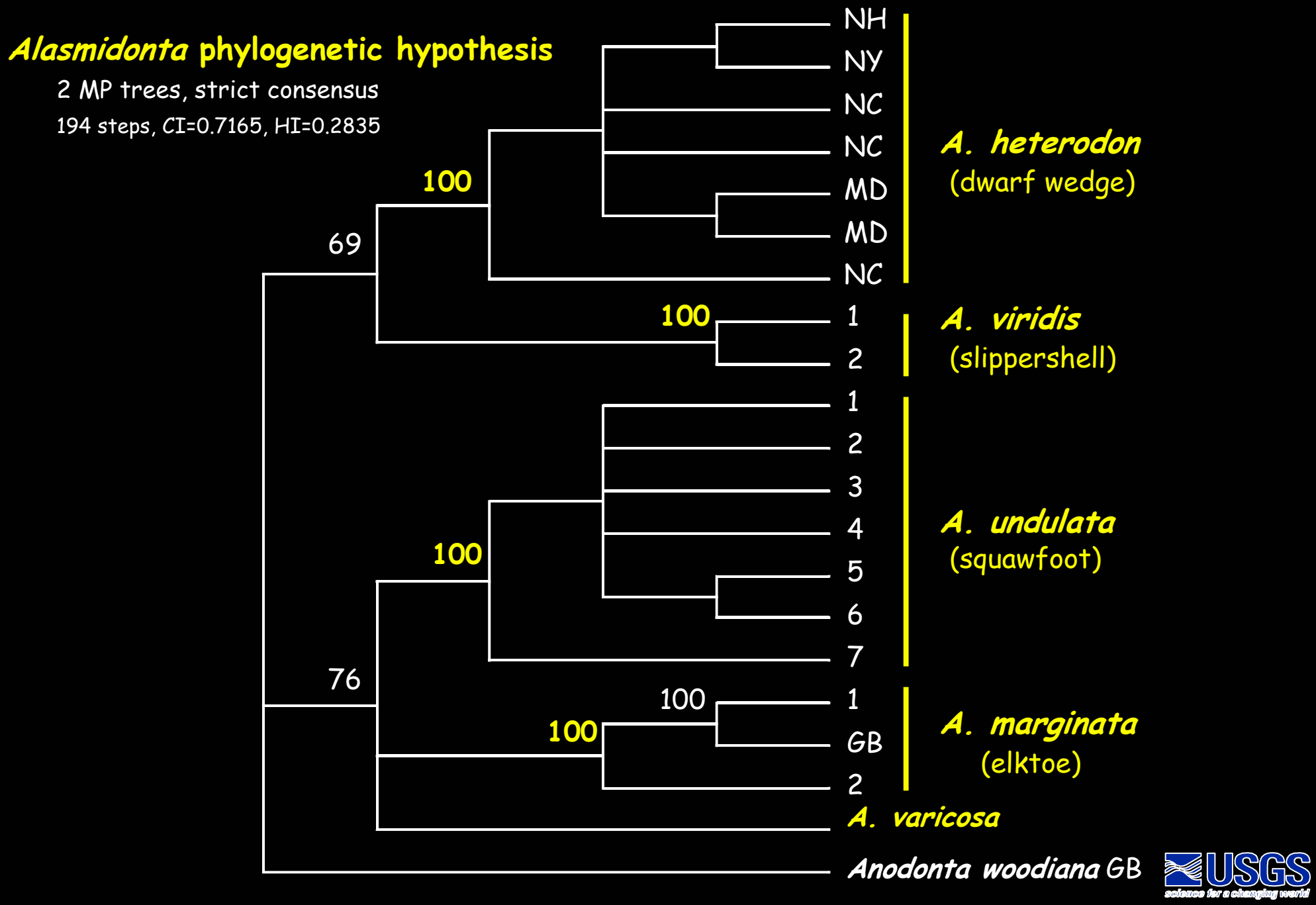
Alasmidonta phylogenetic hypothesis

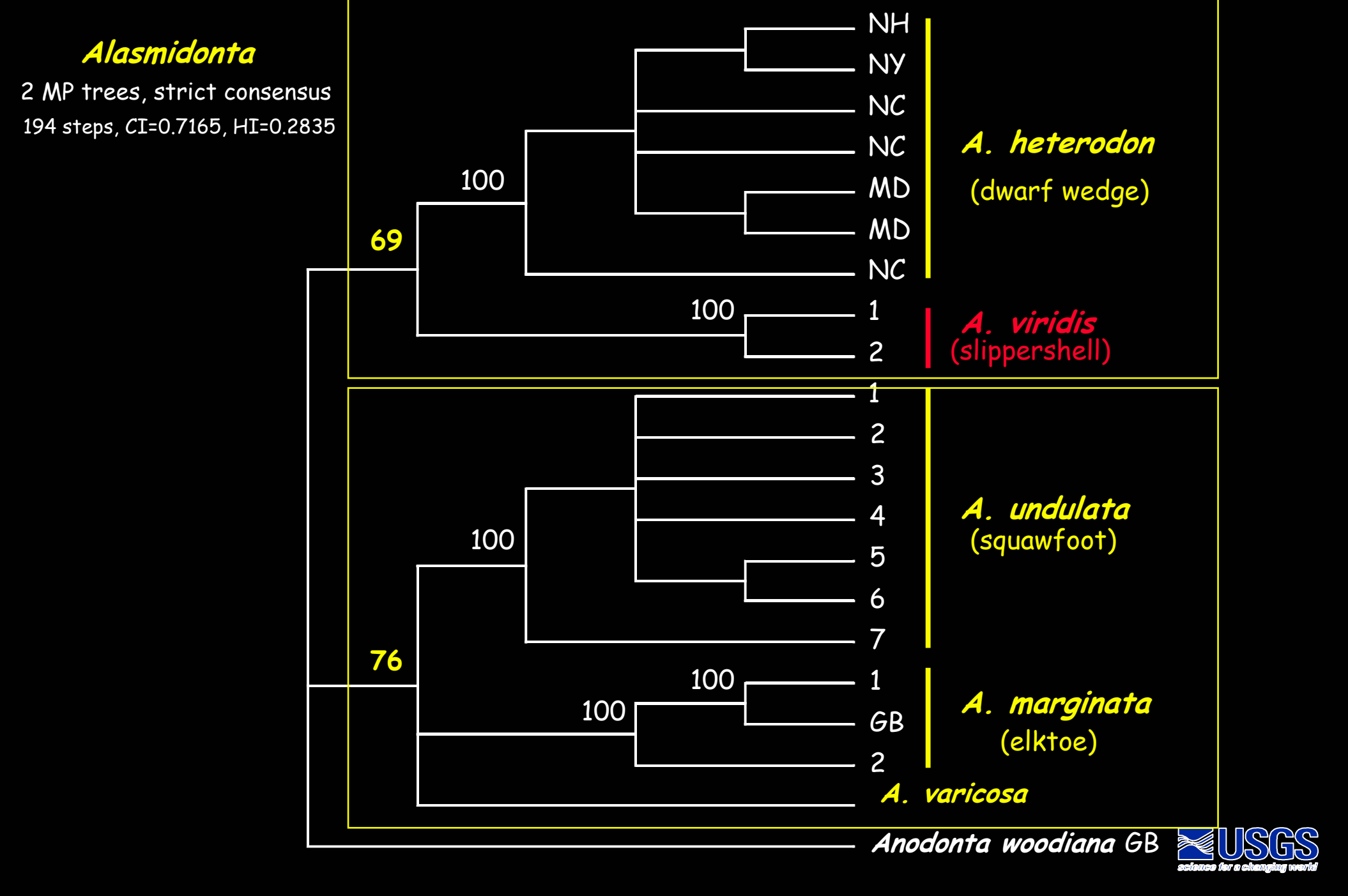
COI: NJ K2P distances



slippershell

USGS
science for a changing world

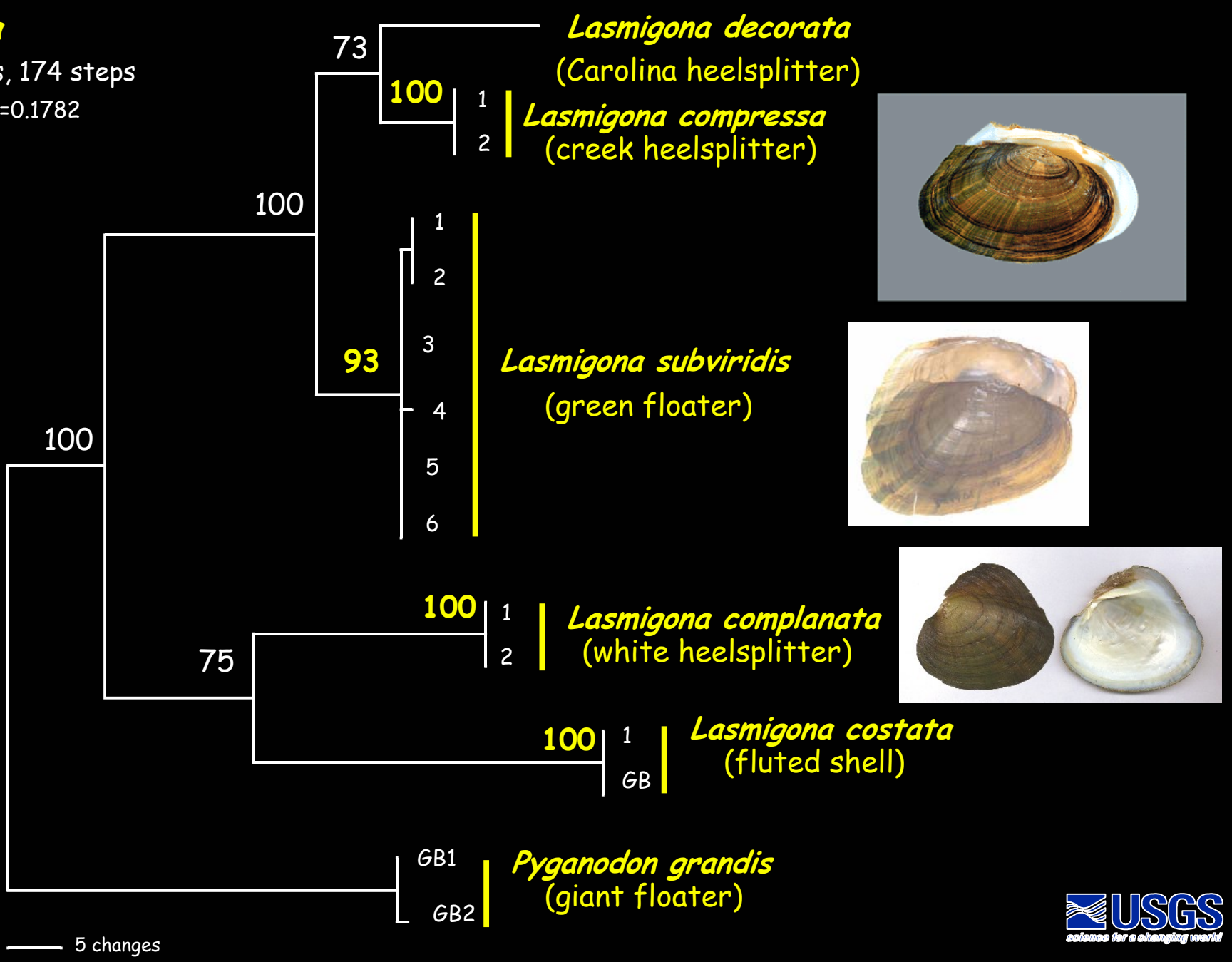

Alasmidonta phylogenetic hypothesis
2 MP trees, strict consensus
194 steps, CI=0.7165, HI=0.2835
NH
NY
NC
NC
MD
MD
NC
1
2
1
2
3
4
5
6
7
1
GB
2
Anodonta woodiana GB
100
69
100
100
76
100
A. heterodon
(dwarf wedge)
A. viridis
(slippershell)
A. undulata
(squawfoot)
A. marginata
(elktoe)
A. varicosa
USGS
science for a changing world


Alasmidonta
2 MP trees, strict consensus
194 steps, CI=0.7165, HI=0.2835
NH
NY
NC
NC
MD
MD
NC
100
69
100
1
2
A. heterodon
(dwarf wedge)
A. viridis
(slippershell)
1
2
3
4
5
6
7
100
76
100
100
1
GB
2
A. undulata
(squawfoot)
A. marginata
(elktoe)
A. varicosa
Anodonta woodiana GB
USGS
science for a changing world

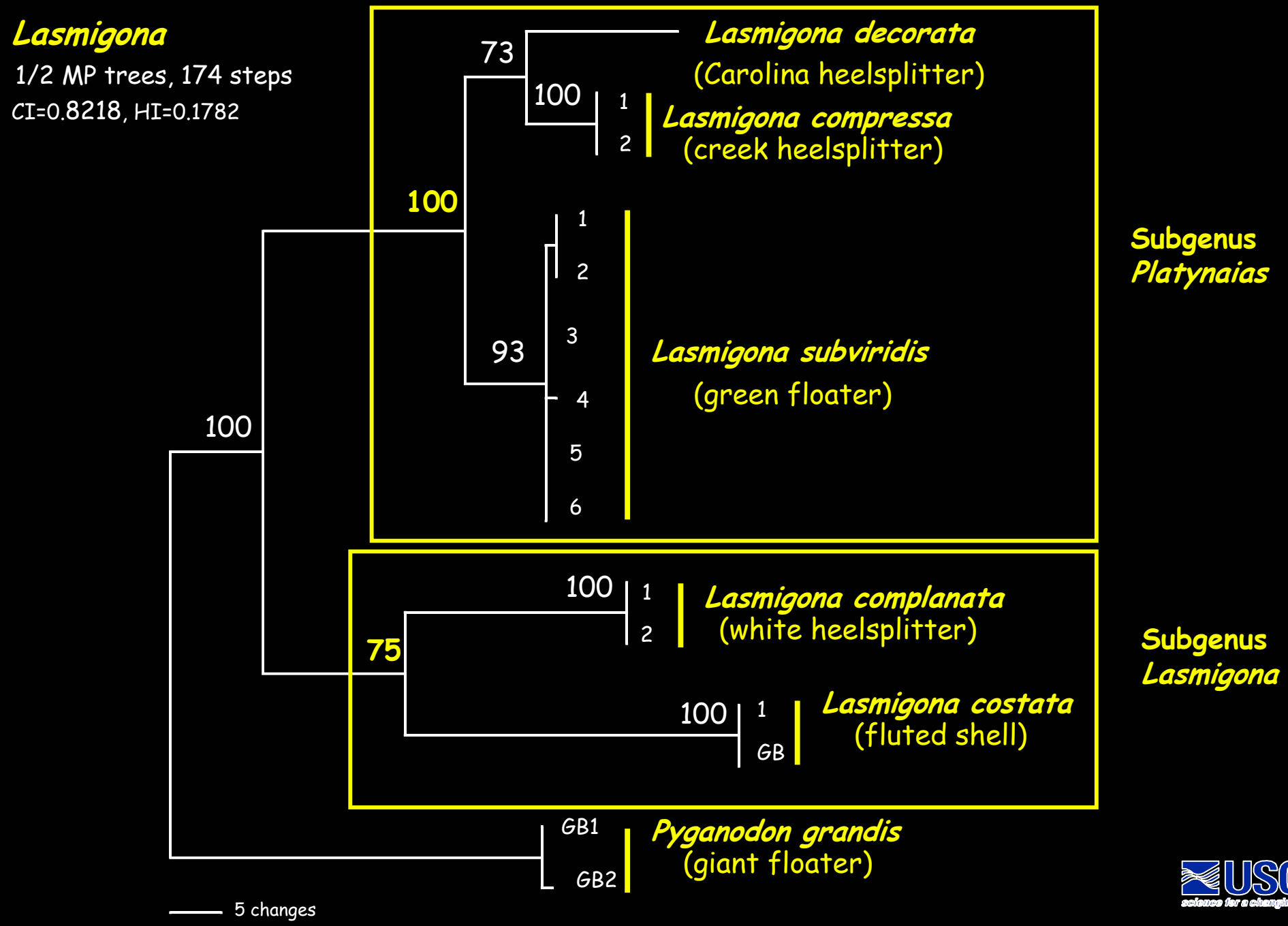
Lasmigona

1/2 MP trees, 174 steps

CI=0.8218, HI=0.1782



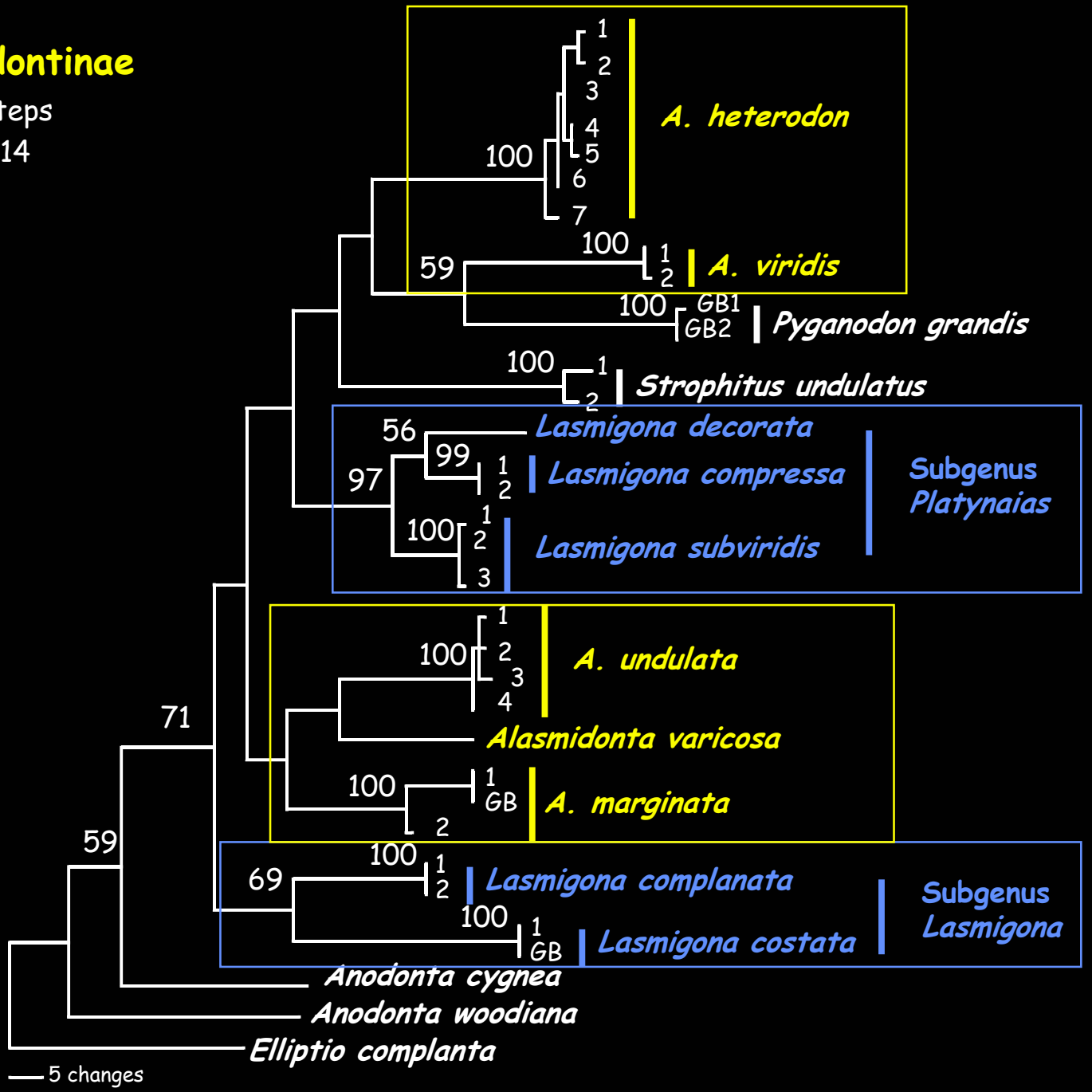
USGS
science for a changing world


Lasmigona
1/2 MP trees, 174 steps
CI=0.8218, HI=0.1782
Lasmigona decorata
(Carolina heelsplitter)
Lasmigona compressa
(creek heelsplitter)
Lasmigona subviridis
(green floater)
Subgenus
Platynaias
Lasmigona complanata
(white heelsplitter)
Lasmigona costata
(fluted shell)
Subgenus
Lasmigona
Pyganodon grandis
(giant floater)
5 changes
USGS
science for a changing world

Subfamily Anodontinae

1/9 MP trees, 514 steps

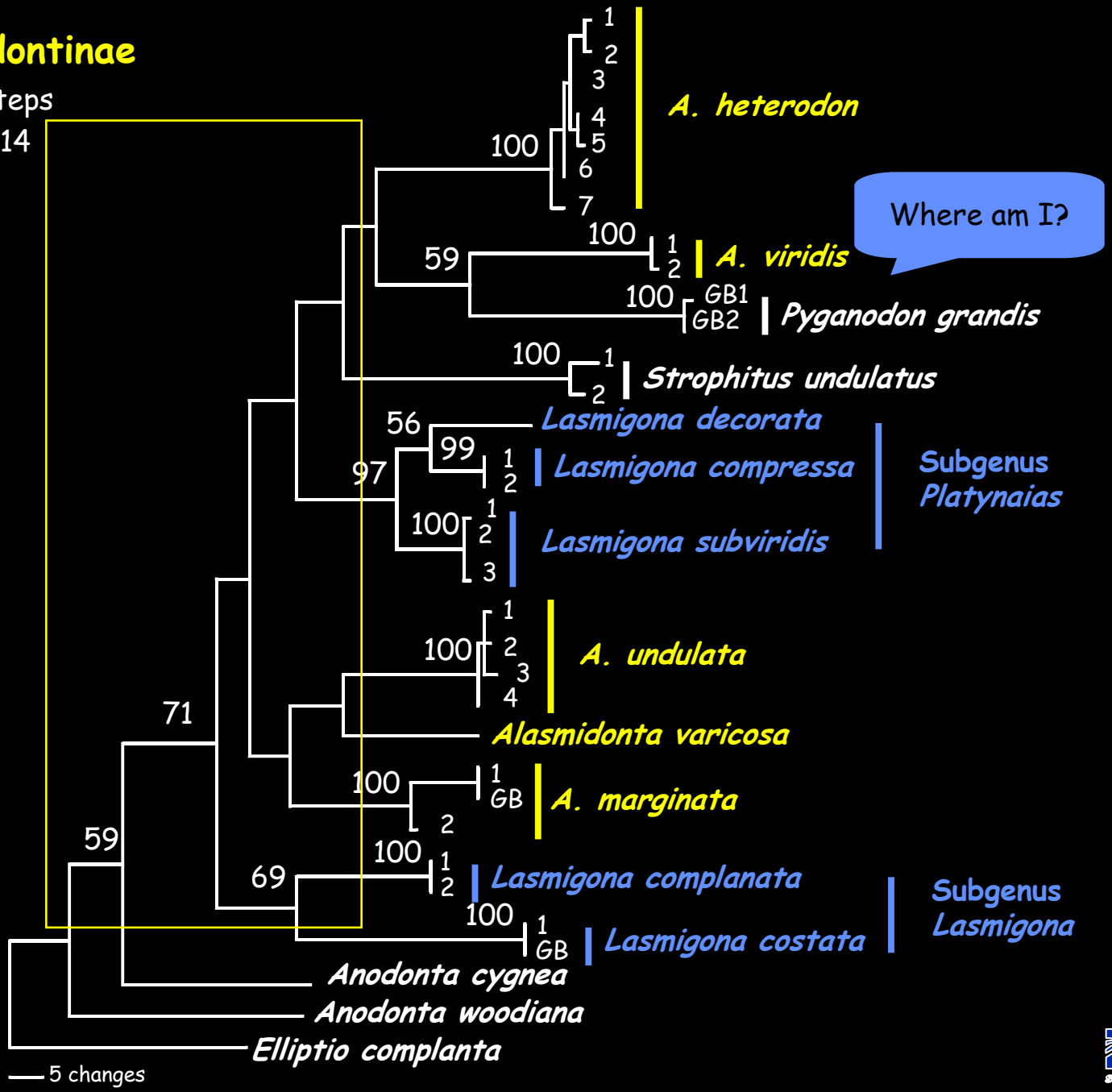
CI=0.4786, HI=0.5214



USGS science for a changing world

Subfamily Anodontinae

1/9 MP trees, 514 steps
CI=0.4786, HI=0.5214



USGS science for a changing world

Conclusions
Subfamily Anodontinae phylogeny

- *Alasmidonta*
  - *(A. heterodon)(A. undulata/ marginata/ varicosa)*
  - *A. viridis* ???
- *Lasmigona*: subgenera recovered
- COI = good tool for species level ID
- Inclusion of many genera = chaos!
  - Not enough signal in data
  - "Prodigious Polyphyly" ???

USGS
science for a changing world

Phylogeography

(Avise 1987)

- Comparison of *phylogenies* of populations or species with their geographic distributions, OR spatial analysis of gene lineages

- Draws from:
  - phylogenetics
  - population genetics
  - biogeography

  } "bridges gap"

- Most common uses –
  - delineate distinct population segments (population subdivision)
  - PRESENT DAY forces- gene flow/effective population sizes
  - HISTORICAL EVENTS- habitat fragmentation and expansion

USGS science for a changing world

Phylogeography

Methods- detecting population structure

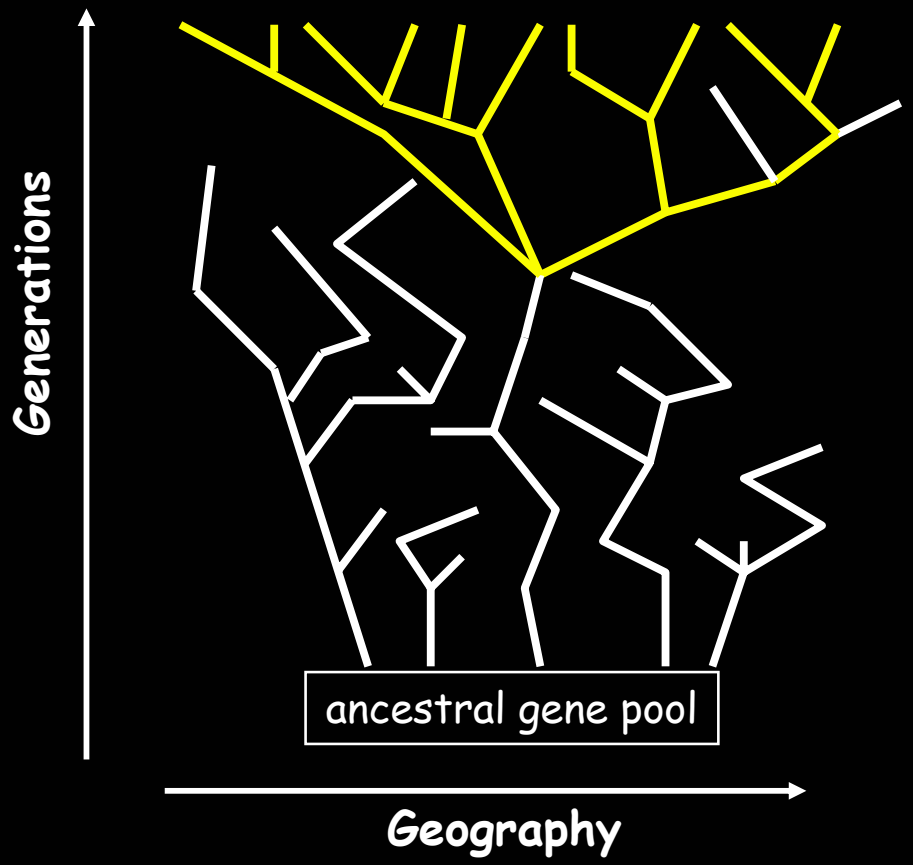
- Geneological-based
  - Phylogenetic analysis
  - Haplotype networks
  - Nested clade analysis
- Frequency-based
  - F-statistics
  - AMOVA

USGS
science for a changing world

Phylogeography

The Coalescent

"model of lineage sorting and genetic drift run backwards in time to common ancestor" (Harding 1996)



- Pathways connecting mothers and daughters
  - **Trace back to 1 female**
- Variance among females in contribution of daughters
  - **Gene tree grows and self-prunes**
- Demographic factors that influence family size important
  - **Chances of lineage survival decrease with population size**

USGS
science for a changing world

Phylogeography

Demography - Phylogeny connections

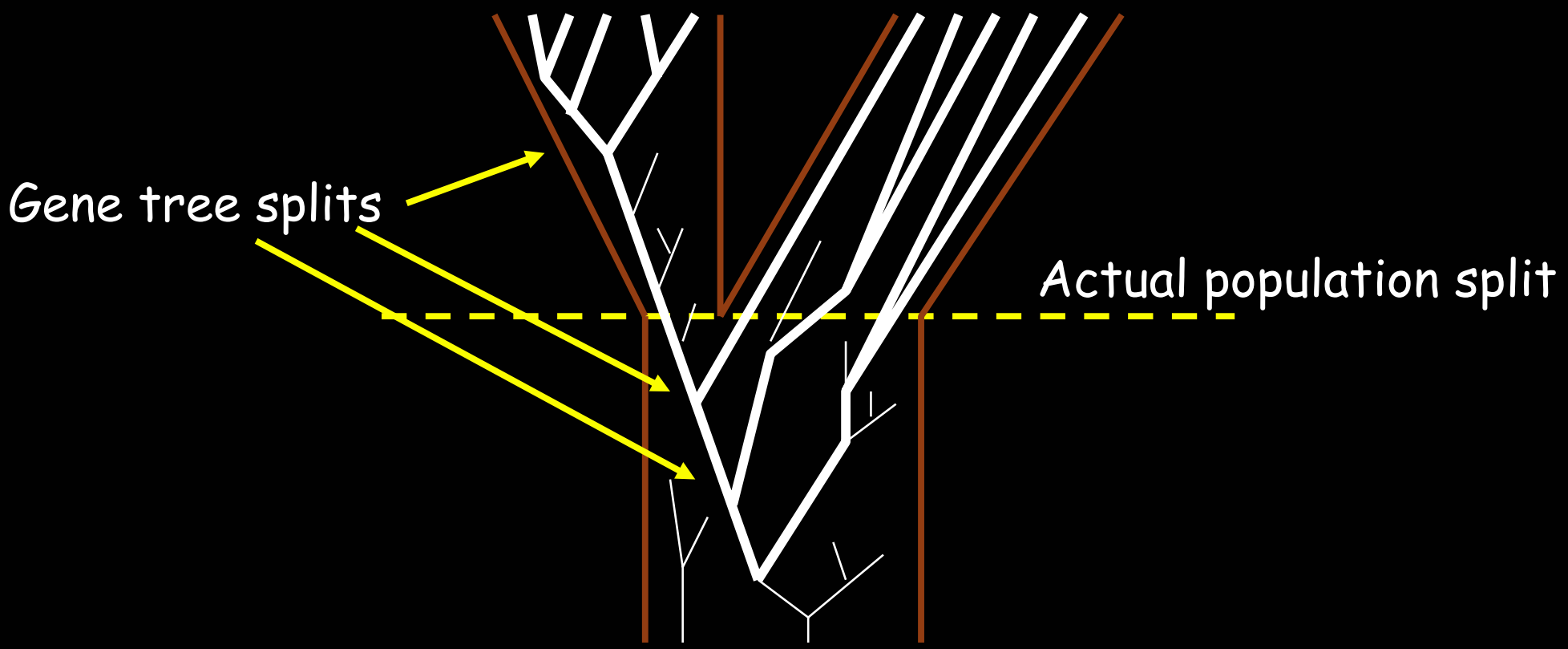
Phylogenetic categories of relationships (Avise 1983)

- Monophyly- barrier concordant with earlier subdivision
  - **female lines trace back to single ancestor**
- Monophyly- barrier concordant with earlier subdivision
  - **female lines trace back to single ancestor**
- Polyphyly- barrier discordant with earlier subdivision
  - **Regions share a lineage**
- Paraphyly- barrier at periphery of range
  - **No unique lineage**

USGS
science for a changing world

"Tree Thinking"

Gene Trees vs. Species Trees



Gene trees can differ from one another and from the species tree

USGS
science for a changing world

Phylogeography

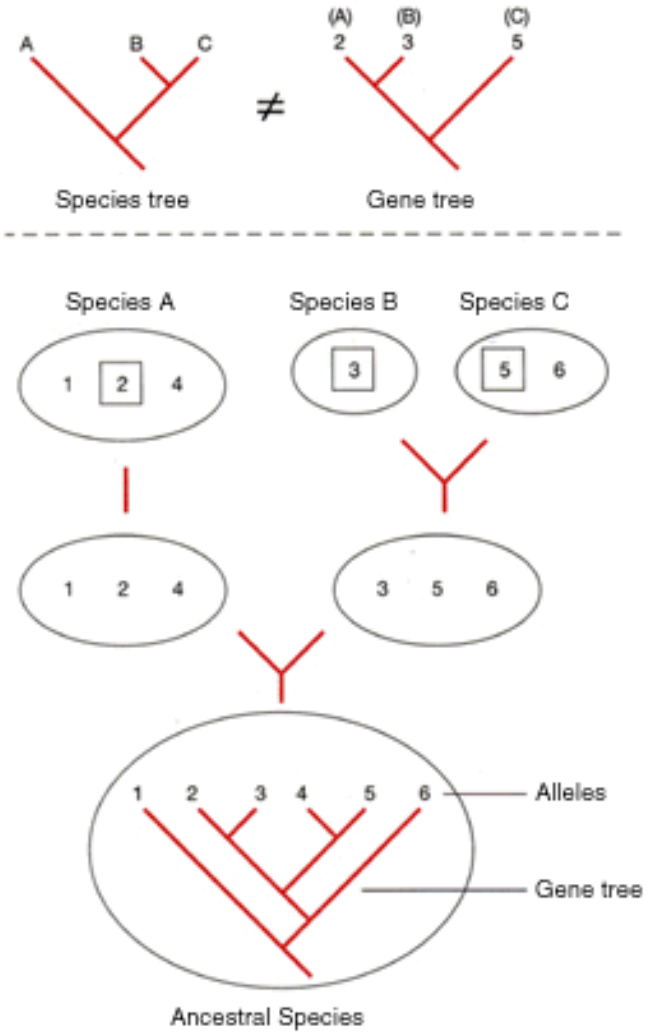
Geneological Concordance (Avise 1996, 2000)

- ESUs:
  - I. Concordance among sequence characters within a gene
  - II. Concordance in major partitions using multiple genes

- Areas of conservation relevance:
  - III. Concordance in geography of gene-tree partitions across multiple species
  - IV. Concordance in gene-tree partitions by geographic provinces

USGS
science for a changing world

Gene Trees vs. Species Trees



Gene Trees vs. Species Trees Read the figure from the bottom to the top. The oval at the bottom represents an ancestral species containing six alleles for a gene. The alleles are related to each other as shown by the gene phylogeny inside the oval: Each allele was derived from its ancestors via a series of mutations.Moving up the figure, a speciation event produces two sister species from the ancestral species. By selection of drift, one species loses alleles 3, 5, and 6, while the other species loses alleles 1, 2, and 4. Moving up the figure, another speciation event occurs, followed by the loss of alleles in its descendant species. We now have: species A, containing alleles 1,2, and 4. 1, 2, and 4; species B, containing allele 3; and species C, containing alleles 5 and 6. Now imagine that we sequence a single allele from each species (boxes) and reconstruct a phylogeny. In the true species tree, species B and C are closest relatives. The gene tree and the species tree show different branching patterns. Note that if we had sampled more extensively from each species and found all six alleles, we would have realized that the gene tree was a misleading guide for estimating the species phylogeny, because the alleles from species A, for example, would not cluster together. After Ruvolo (1994).

USGS
science for a changing world

Species Concepts and Criteria

USGS
science for a changing world

Population Genetics and Molecular Systematics

- Different biological disciplines
- Genetic data have enabled an interface of sorts between the two—which is good, but potentially dangerous when the parameters of the study in question are not explicit
- Genetic data are not magic

USGS
science for a changing world

Population Genetics and Molecular Systematics

- While the techniques of genetics have opened doors (or windows) between fields, at times fundamental concepts have had to be recast; the wheel has been reinvented (in my opinion unnecessarily) on a number of occasions
- Nevertheless these debates have enabled a clearer understanding of the complementarity of systematics and population genetics; it's not as though one were superior: each is best equipped to deal with certain kinds of questions

USGS
science for a changing world

Population Genetics, Molecular Systematics, and Species

- "The species problem" has not, nor will it ever go away
- Species concepts have been viewed as a nexus for synthesizing biology and as distraction from and obstacle to scientific progress
- Regardless, understanding the perspective of any given study towards species and speciation is critical to interpretation

USGS
science for a changing world

Is it important? Yes.

- Is it ok that people disagree?
  Absolutely.

USGS
science for a changing world

General Background

- Various, often conflicting approaches to "the" species problem, confounded by the conflation of different kinds of genetic data

USGS
science for a changing world

General Background

- Various, often conflicting approaches to "the" species problem , confounded by the conflation of different kinds of genetic data
- Literature devoted to the species problem has not abated

USGS
science for a changing world

ELSEVIER

Delimiting species: a Renaissance issue in systematic biology

Jack W. Sites Jr and Jonathon C. Marshall

Department of Integrative Biology and M.L. Bean Life Science Museum, Brigham Young University, Provo, UT 84602-5181, USA

ELSEVIER

Understanding and confronting species uncertainty in biology and conservation

Jody Hey[1], Robin S. Waples[2], Michael L. Arnold[3], Roger K. Butlin[4] and Richard G. Harrison[5]

[1]Department of Genetics, Rutgers University, Piscataway, NJ 08854, USA
[2]National Marine Fisheries Service, Northwest Fisheries Science Center, Seattle, WA 98112, USA
[3]Department of Genetics, University of Georgia, Athens, GA 30602, USA
[4]Centre for Biodiversity and Conservation, School of Biology, The University of Leeds, Leeds, UK LS2 9JT
[5]Department of Ecology and Evolutionary Biology, Cornell University, Ithaca, NY 14853, USA

USGS science for a changing world

General Background

- Various, often conflicting approaches to "the" species problem , confounded by the conflation of different kinds of genetic data
- Literature devoted to the species problem has not abated
- Legislation has not kept up with current debates and issues (as if any of us could)

USGS
science for a changing world

Words, words, words

- Arguments over definitions lead nowhere…by definition…

USGS
science for a changing world

Species, speciation, units, and continua

- Woodger's dilemma: Species boundaries and the biological species concept

USGS
science for a changing world

Species, speciation, units, and continua

- Woodger's dilemma: Species boundaries and the biological species concept

- Species concepts and criteria: What are they supposed to do?

USGS
science for a changing world

Species, speciation, units, and continua

- Woodger's dilemma: Species boundaries and the biological species concept
- Species concepts and criteria: What are they supposed to do?
- Do we need to regard species as biologically equivalent in terms of their origin and/or maintenance?

USGS
science for a changing world

Species, speciation, units, and continua

Limits of the biological species concept

- Confined to extant biparental sexually reproducing organisms
- Admits a continuum in which the partitions may not reflect history
- Reproductive isolation (as "speciation") may occur late in the "ontogeny" of species concepts

USGS
science for a changing world

Species concepts and criteria

- "Biological"
- Cohesion
- Phylogenetic
  - I
  - II
  - III…
- "Cladistic"
- Morphological

USGS
science for a changing world

Species

Axes of Understanding: Ontological versus Graphic Features of Various Species Concepts and Criteria

USGS
science for a changing world

Ontological Features of Various Species Concepts, Including Versions of Phylogenetic Species

Species concept	Author/proponent	Criterion	Emphasis	Ontological features				
				Criteria general vs context-dependent	Mechanistic vs theory-neutral	Pluralistic vs monistic	Individuals vs classes	Discrete vs continuous
BSC	Mayr, Dobzhansky, Avise	Potential interbreeding	Maintenance	Context-dependent	Mechanistic	Monistic	Classes	Continuous
EvSC	Simpson, Wiley	Lineage cohesion	Maintenance	General	Theory-neutral	Monistic	Classes	Continuous
ISC	Hennig, Brundin, Ridley	Lineage cohesion	Maintenance	General	Theory-neutral	Monistic	Individuals	Discrete
EcSC	Van Valen	Adaptive peaks	Maintenance	General	Mechanistic	Monistic	"Individualistic classes"	Continuous
RSC	Paterson	Isolating mechanisms	Maintenance	Context-dependent	Mechanistic	Monistic	—	Discrete
CSC	Templeton	Reproductive cohesion	Maintenance	Context-dependent	Mechanistic	Monistic	Individuals	Discrete
PSC	Cracraft, Nixon, Davis, Wheeler, Luckow	Diagnostic characters	Origin	General	Theory-neutral	Monistic	Individuals	Discrete
PSC	Brower	Haplotype networks	Origin	General	Theory-neutral	Monistic	Individuals	Discrete
ASC	de Queiroz, Donoghue, Olmstead, Hill and Crane	Monophyly	Origin	General	Mechanistic	Pluralistic	Individuals	Continuous
GSC	Baum, Donoghue, Shaw	Monophyly	Origin	General	Mechanistic	Pluralistic	Individuals	Continuous

Note. BSC, biological species concept; EvSC, evolutionory species concept; ISC, internodal species concept; EcSC, ecological species concept; RSC, recognition species concept; CSC, cohesion species concept; PSC, phylogenetic species concept; ASC, actapomorphic species concept; GSC, geneological species concept.

Cladistic Features and Implications of Various Species Concepts, Including Versions of Phylogenetic Species

Species concept	Author/proponent	Temporal component	Cladistic Corollaries		Cladogram graphic	Change in both descendant species
			Perspective	Type of change		
BSC	Mayr, Dobzhansky, Avise	Unidimensional	Current	Anagenetic,	None	Required
EvSC	Simpson, Wiley	Time-extended	Prospective	Anagenetic, cladoge-netic	Internodal	Required
ISC	Hennig, Brundin, Ridley	Time-extended	Retrospective	Anagenetic, cladoge-netic	Internodal	Not required
EsSC	Van Valen	Unidimensional	Current	Anagenetic,	None	Required
RSC	Paterson	Unidimensional	Current	Anagenetic,	None	Not required
CSC	Templeton	Unidimensional	Current	Anagenetic,	None	Required
PSC	Cracraft, Nixon, Davis, Wheeler, Luckow	Unidimensional	Retrospective	Cladogenetic	Terminal	Not required
PSC	Brower	Unidimensional	Retrospective	Cladogenetic	Network-based	Not required
ASC	de Queiroz, Donoghue, Olmstead, Hill, and Crane	Unidimensional	Retrospective	Anagenetic, cladoge-netic	Tree-based ("mono-phyly')	Required
GSC	Baum, Donoghue, Shaw	Time-extended	Retrospective	Anagenetic, cladoge-netic	Tree-based ("mono-phyly")	Required

Species concepts and conservation

- Species, subspecies, ESUs, etc
- Recognition, history, process, and potential

USGS
science for a changing world

Magic Numbers?

- Percent divergence as a criterion for species delineation is necessarily arbitrary

USGS
science for a changing world

Diagnosing Units of Conservation Management

ALFRIED P. VOGLER
ROB DESALLE
American Museum of Natural History
Central Park West at 79th Street
New York, NY 10024, U.S.A.

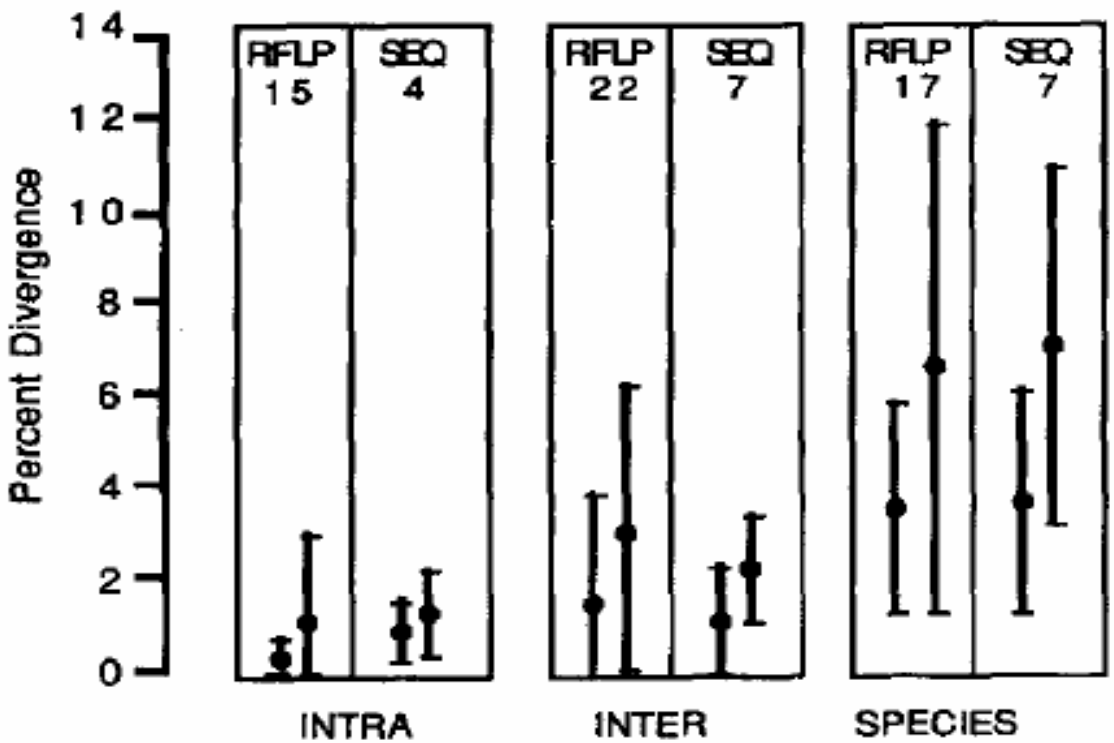


Figure 1. Average minimum (left in each panel) and average maximum (right in each panel) sequence divergence in mtDNA at various hierarchical taxonomic levels. The level of sequence divergence was taken from the literature from restriction fragment length polymorphism (RFLP) or from DNA sequencing (SEQ) studies. The data were split into three taxonomic levels as designated by the authors of the various publications: within populations (INTRA), between populations (INTER), and between "closely related species" (SPECIES). Variances of the mean minimum and maximum sequence divergence are indicated by bars. Both variances are given in the figure. The data were taken from a selection of papers published before the middle of 1991 (see Appendix). The total number of studies included in each of the comparisons is given at the top of each panel. The full data set used in this figure is available from the authors upon request.

Conservation Biology, Pages 354–363
Volume 8, No. 2, June 1994

Tree-based conceptions versus Character-based Criteria

USGS
science for a changing world

Example: crystal darter

Crystallaria asprella

- Once widely distributed in Eastern US
- Severe population decline- habitat degradation
- Exists in isolated populations
  - 6 Southern states: MS and Gulf drainages
  - Recently discovered Elk River, WV
- Assess degree of connectivity between populations

(Morrison, Lemarie, Wood, and King, in press)

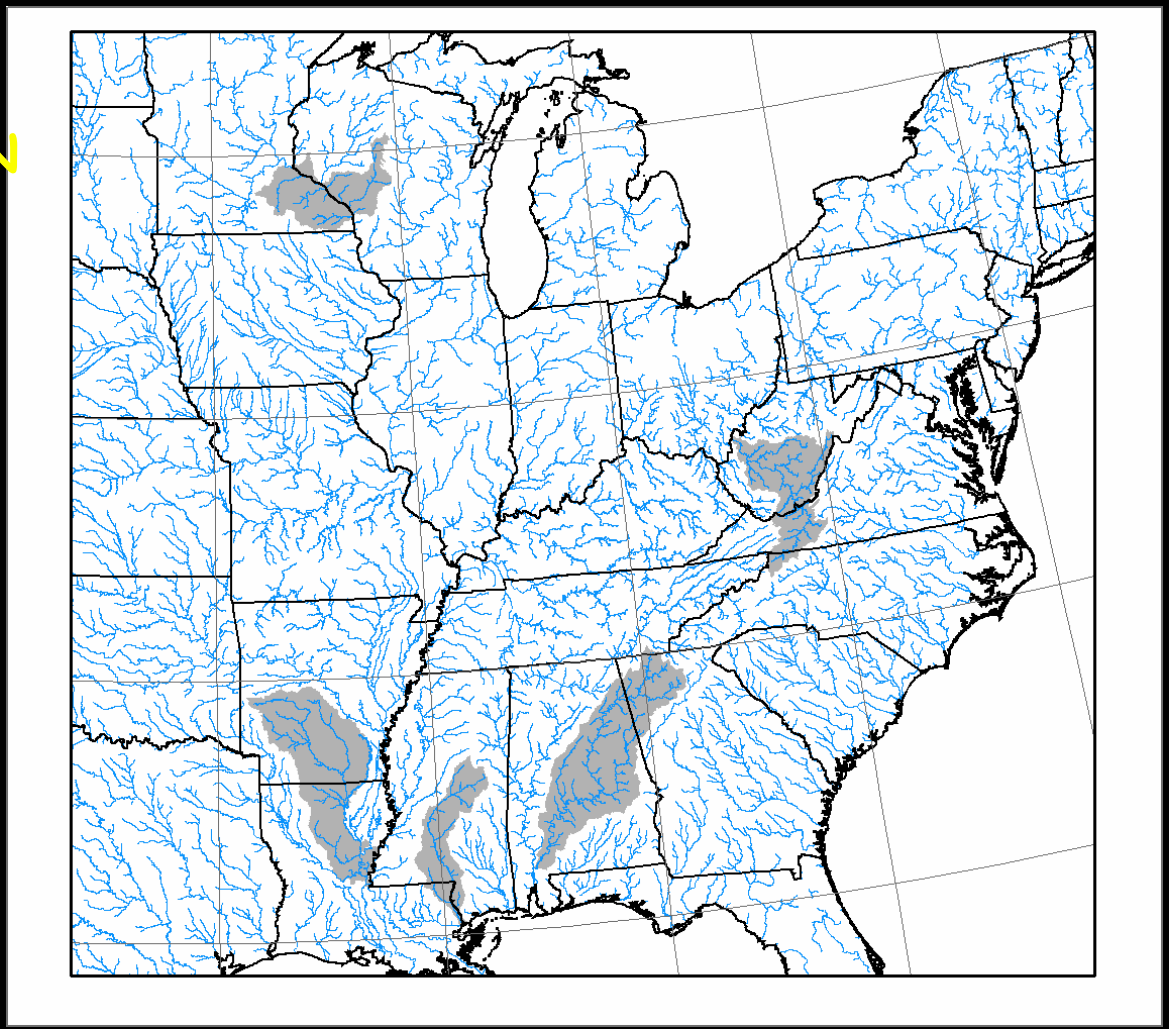
USGS
science for a changing world

Example: crystal darter

Crystallaria asprella

Zumbro River, MN

Saline River, AR



Elk River, WV

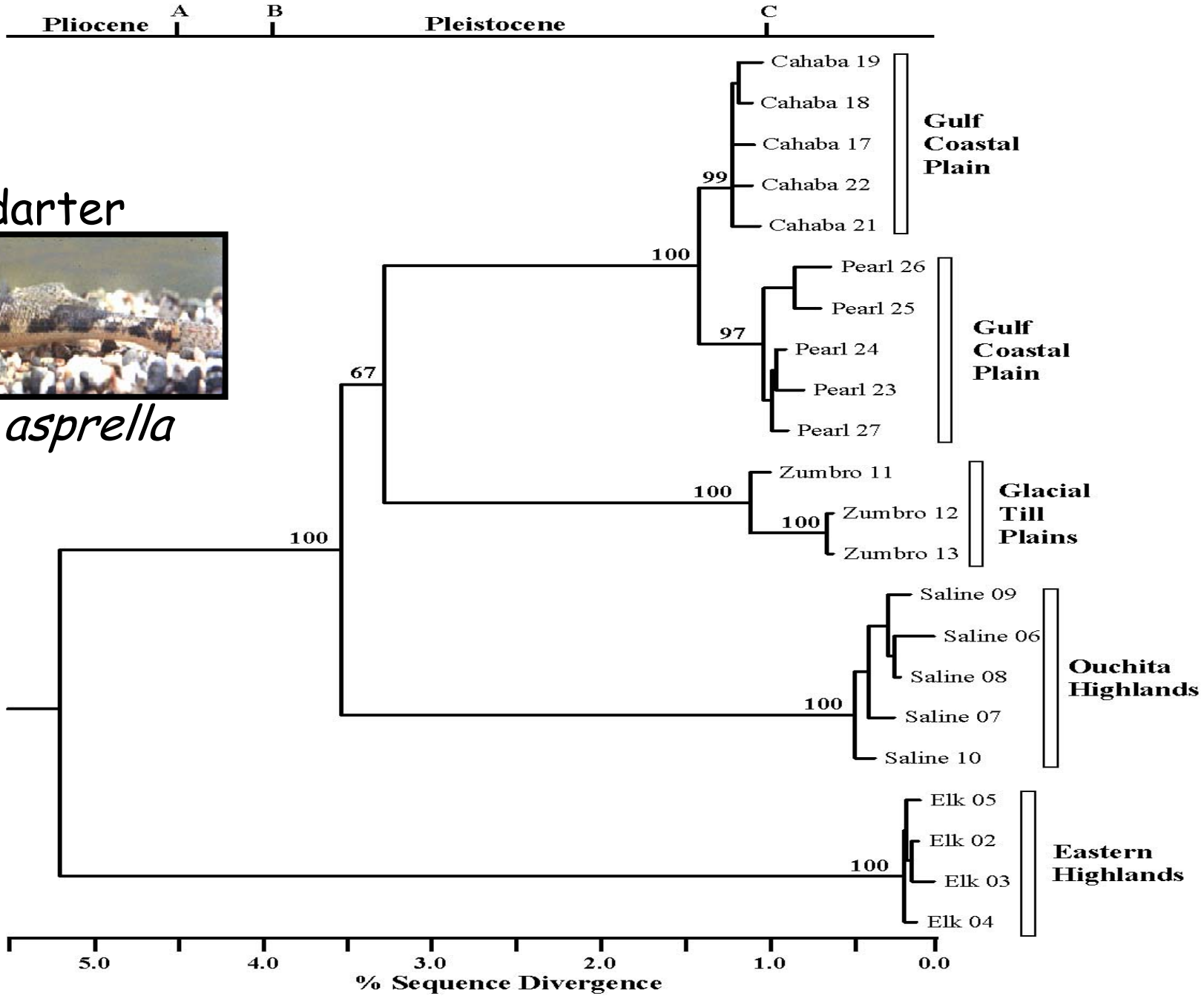
Pearl River, MS

Cahaba River, AL

USGS
science for a changing world

Crystal darter

Crystallaria asprella



Accurate Taxonomy in Conservation Biology

Does it really matter?

- May not be recognizing extent or limits of diversity
  - Genetically distinct (cryptic) species unrecognized?
  - Legal protection for abundant species?
  - Overlooking sources of genetic diversity to help impoverished populations
- Confusion caused by hybrids
- Forensics- identifying poached species

USGS science for a changing world

Trees as Networks

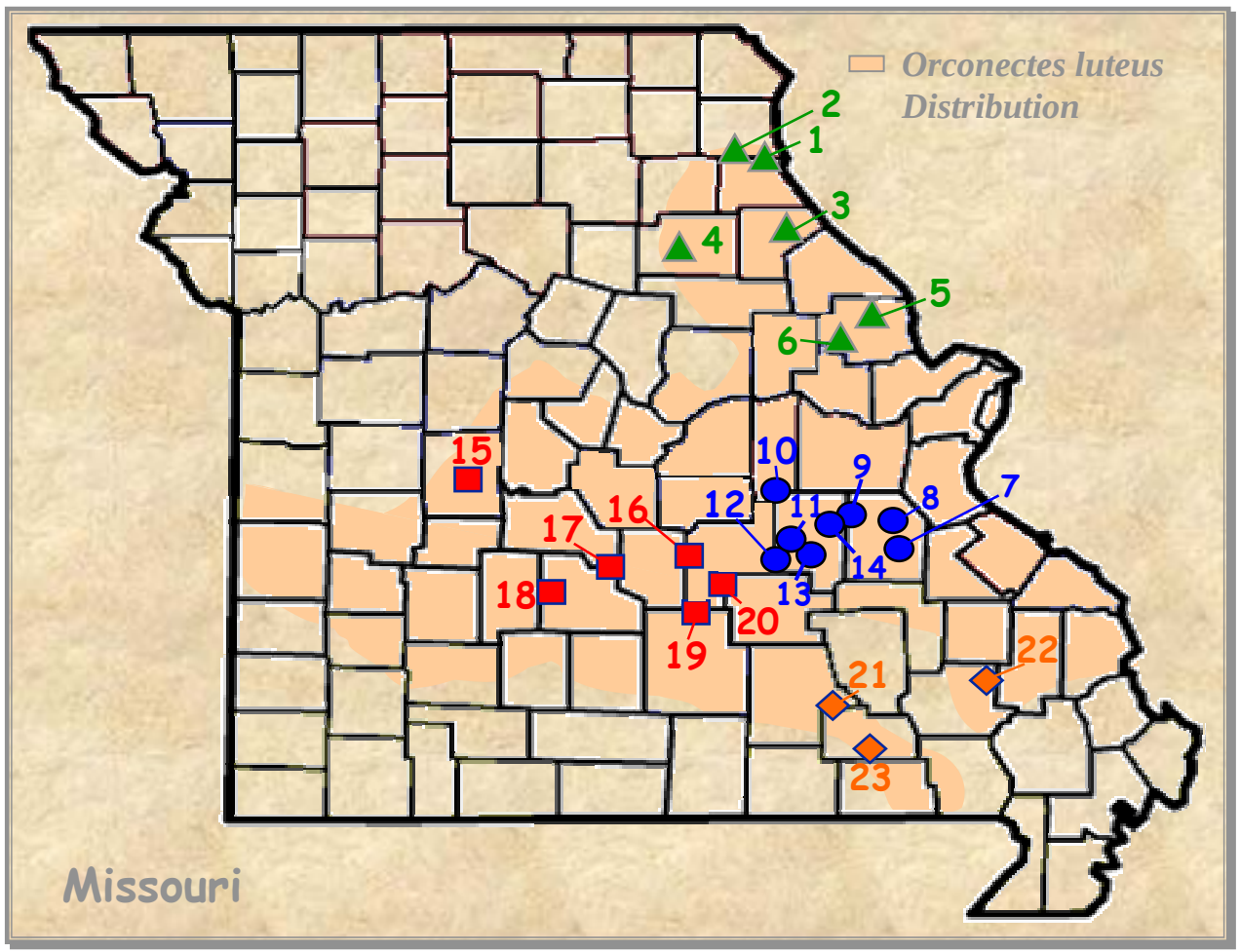
- The interface of Systematics and Population Genetics
- Network incorporates reality of population genetics
  - Reticulate relationships
  - Hybridization
  - Back mutation
  - Recombination

USGS
science for a changing world

Population Genetics: *Orconectes luteus*

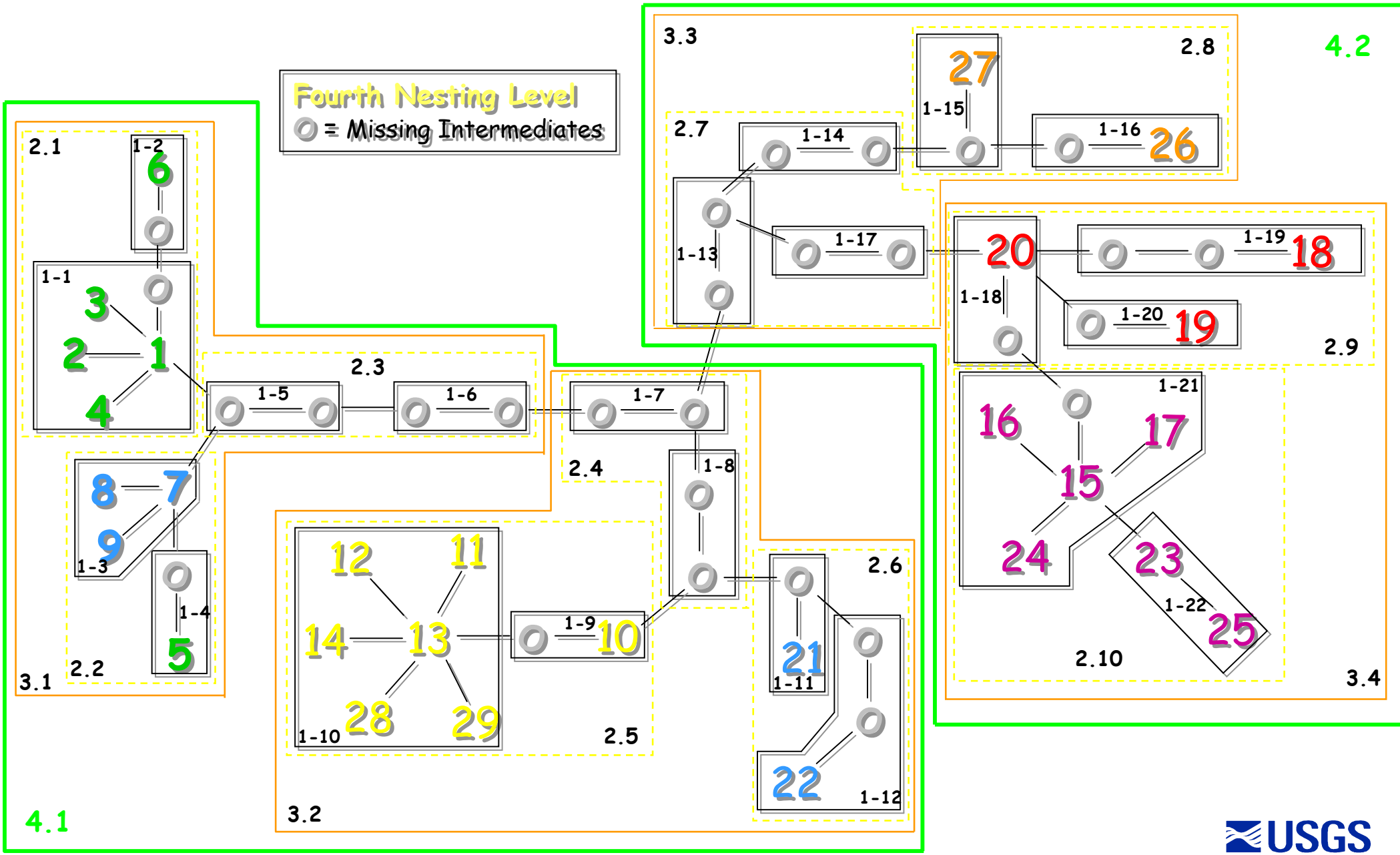
USGS
science for a changing world

Sample Localities for *Orconectes luteus*



▲ = Northeastern Missouri Populations
● = Meramec River Populations
■ = Missouri River Populations
◆ = Current & St. Francis River Pops.

USGS
science for a changing world


Fourth Nesting Level
= Missing Intermediates
4.1
4.2
3.1
3.2
3.3
3.4
2.1
2.2
2.3
2.4
2.5
2.6
2.7
2.8
2.9
2.10
1-1
1-2
1-3
1-4
1-5
1-6
1-7
1-8
1-9
1-10
1-11
1-12
1-13
1-14
1-15
1-16
1-17
1-18
1-19
1-20
1-21
1-22
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
USGS

Conclusions

- High estimate of FST from 16s Data (FST=0.937)
  - High level of Population Structuring

- Nested Analysis Indicates:
  - Past fragmentation involved with higher-level (older) clades.
  - Restricted gene-flow (Isolation By Distance) between Big River and Merimac River Populations.
  - Contiguous Range Expansion for Current River, Little Piney, and Big Piney Populations.

- Three Distinct ESUs lacking historical and recent genetic and ecological (color morph & geography) exchangeability - distinct species! (with further genetic subdivision within these groups)

USGS

Optimality Criteria

- Maximum Parsimony (minimize character change over a tree - tree length)
- Maximum Likelihood (maximize the likelihood)
- Minimum Evolution (obtain all trees within a certain genetic distance)
- NOT neighbor-joining, etc. - algorithms

USGS
science for a changing world

Models for all!

- MP - Model is implicit (1:1) or explicit (weighting matrix)
- ML - Model is explicit
- ME - Model is explicit

- How do you choose a model?

USGS
science for a changing world

Hypothesis Testing Hierarchy

(Huelsenbeck & Crandall 1997)

Equal Base Frequencies

JC69 vs. F81

Transition Rate Equals Transversion Rate

JC69 vs. K2P or F81 vs. HKY85

Transition Rates and Transversion Rates Are Equal

Rates Equal Among Sites

USGS

The Phylogenetic Problem

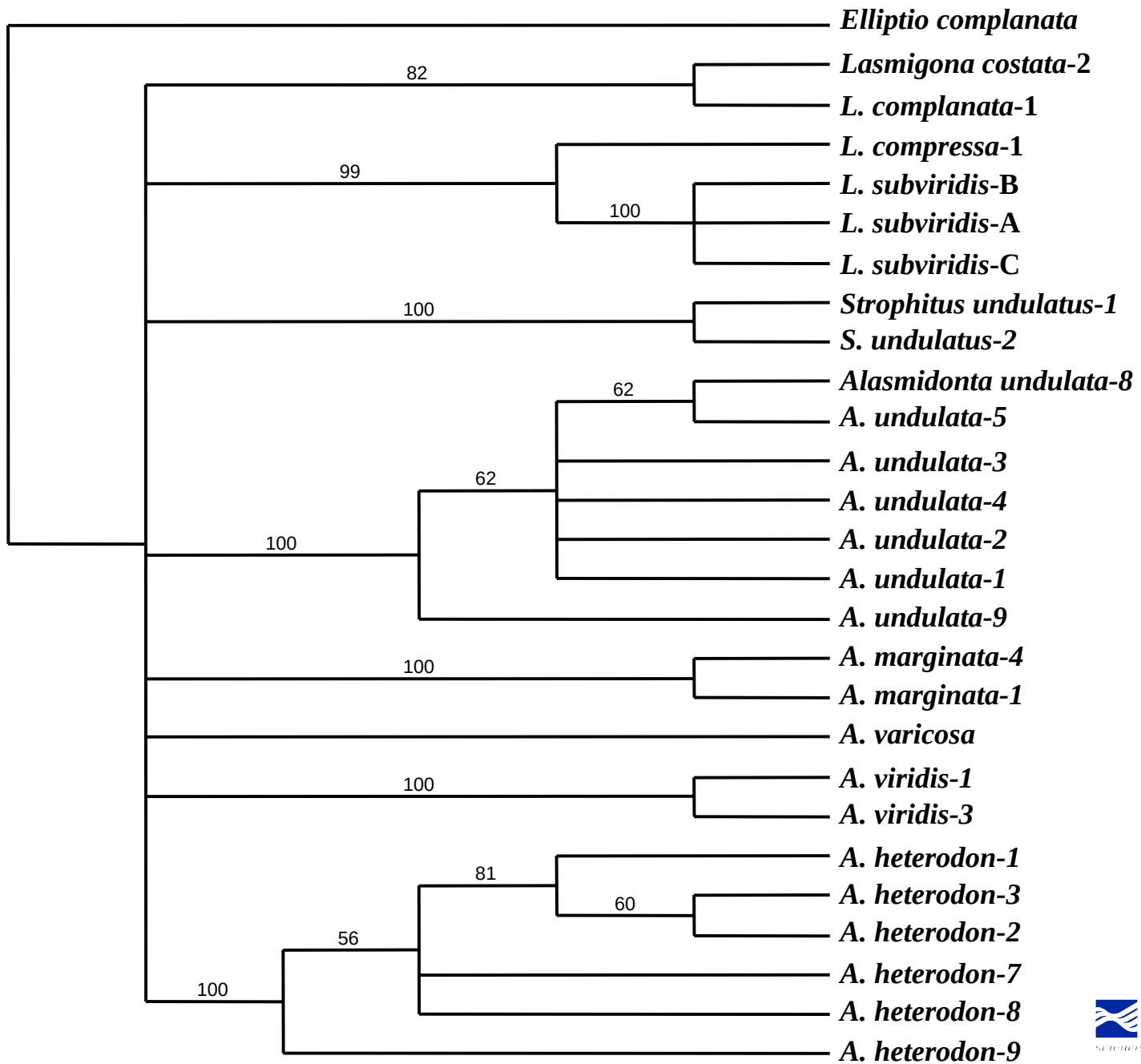
$$B(T) = \prod_{i=3}^{T} (2i - 5)$$

Number of Seqs	Number of Trees
10	$2x10^{6}$
100	$2x10^{182}$
1,000	$2x10^{2,860}$
10,000	$8x10^{38,658}$
100,000	$1x10^{486,663}$
1,000,000	$1x10^{5,866,723}$

USGS
science for a changing world

Alasmidonta

COI
MP Tree



Now you have a tree - what's next?

- Confidence in a tree - bootstrap
- Hypothesis testing with a tree (likelihood approaches)
- Testing for correlations with phylogenetic diversity

USGS
science for a changing world

Aims of systematics

- To describe and arrange our observations and understanding of life on Earth in an empirical, efficient, and recoverable manner

USGS
science for a changing world

Aims of systematics

- To describe and arrange our observations and understanding of life on Earth in an empirical, efficient, and recoverable manner
- To provide an information retrieval system proscribed in such a way that alterations and new discoveries can be accommodated and tracked through one or more scientific codes (e.g. ICZN)

USGS
science for a changing world

Aims of systematics

- To enable the furtherance of all biological understanding, ecological, behavioral, biogeographic, and otherwise, through an information system that facilitates empirical testing and analysis of evolutionary history. That is, by constructing a lens through which evolution itself may be explored

USGS
science for a changing world

Systematics does not equal taxonomy

Taxonomy does not equal nomenclature

- Taxonomy is an arm of systematics devoted to description, nomenclature, and classification

USGS
science for a changing world

History

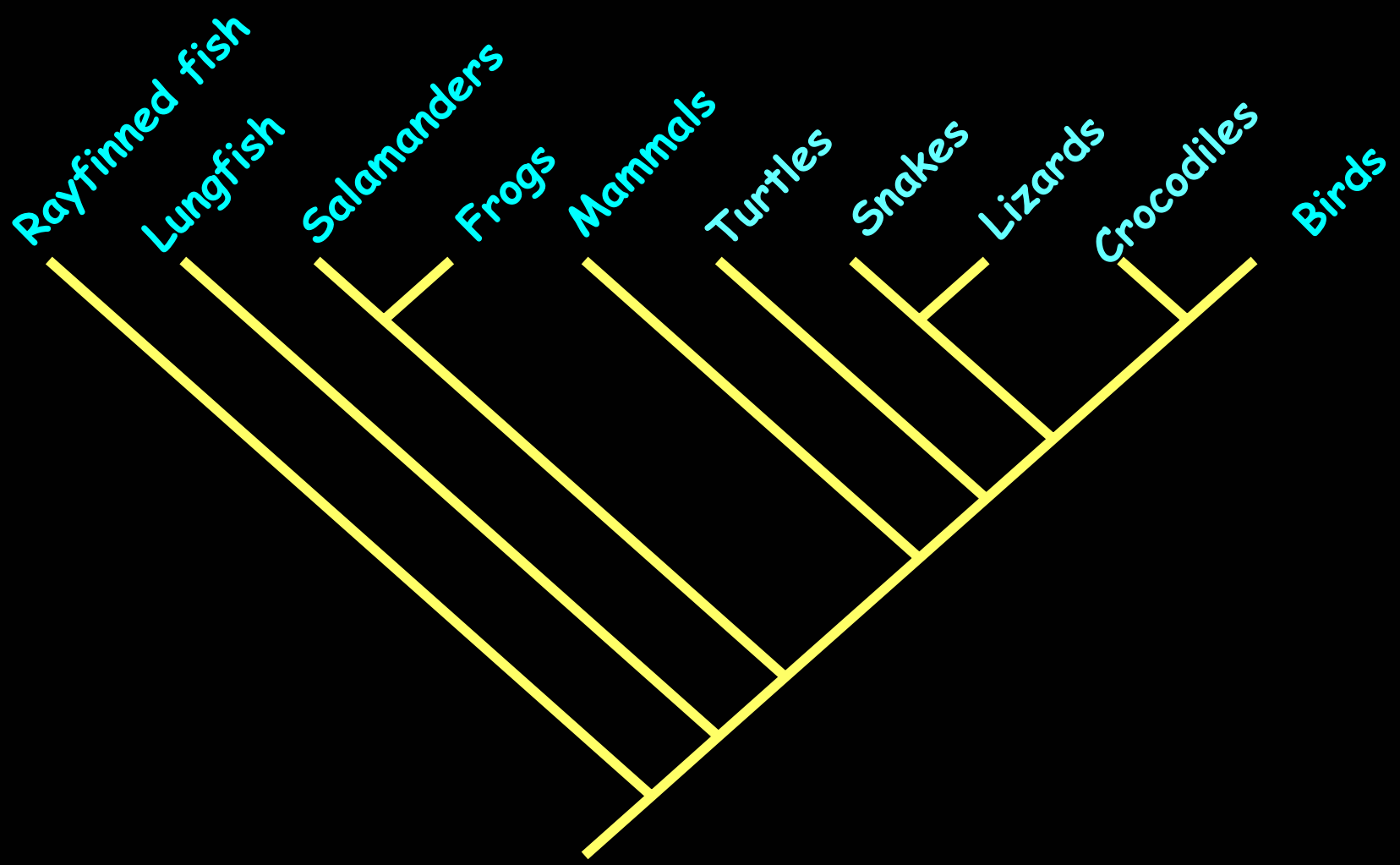
- Early debates surrounded "schools" of systematics: Evolutionary taxonomists, "numerical taxonomists" (who embraced phenetic methods), and phylogenetic systematists (cladists)

USGS
science for a changing world

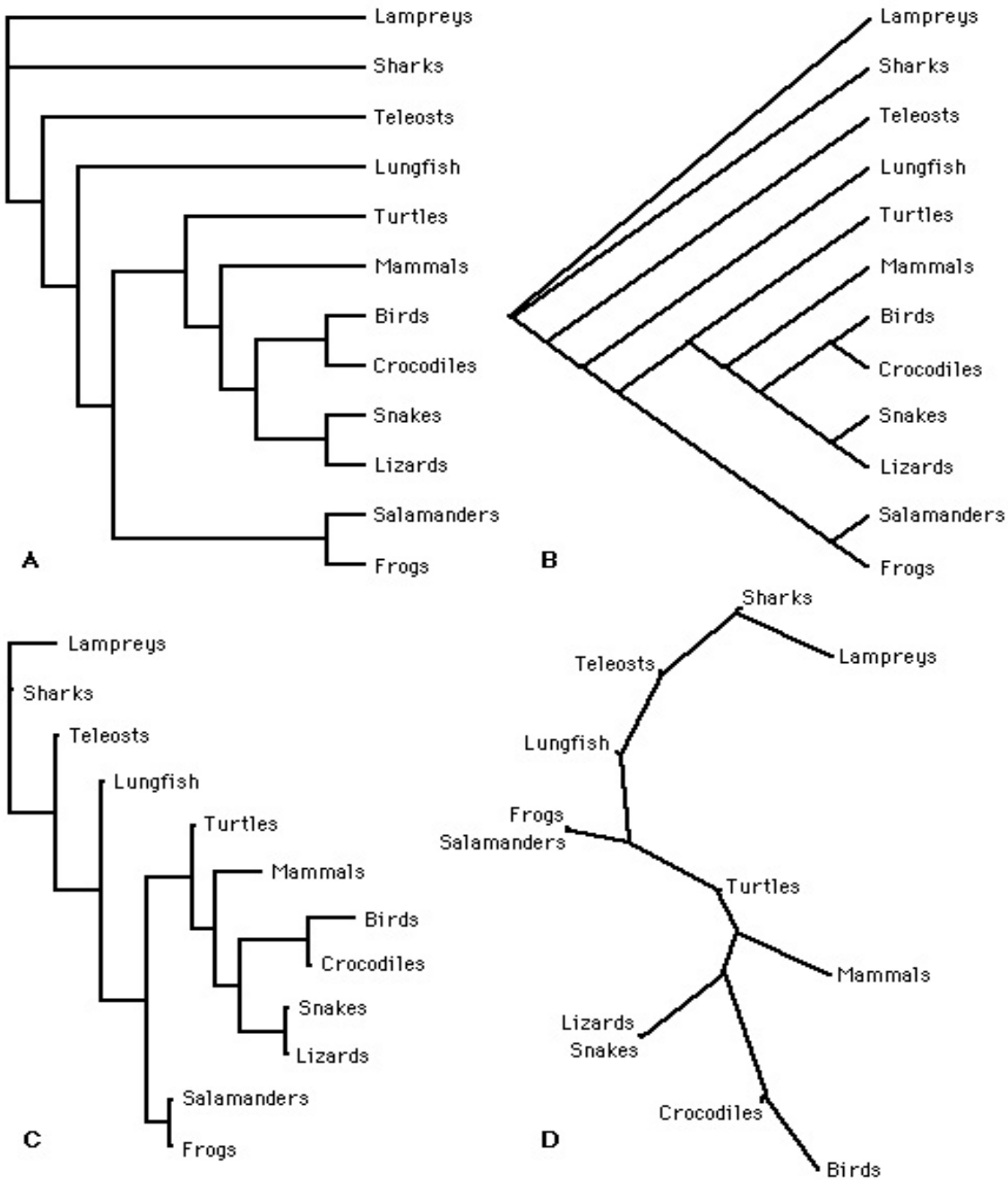
Phylogenetic Systematics (Hennig, 1966)

- Incorporates taxonomy
- Involves quantification, character analysis, and hypothesis testing

USGS
science for a changing world



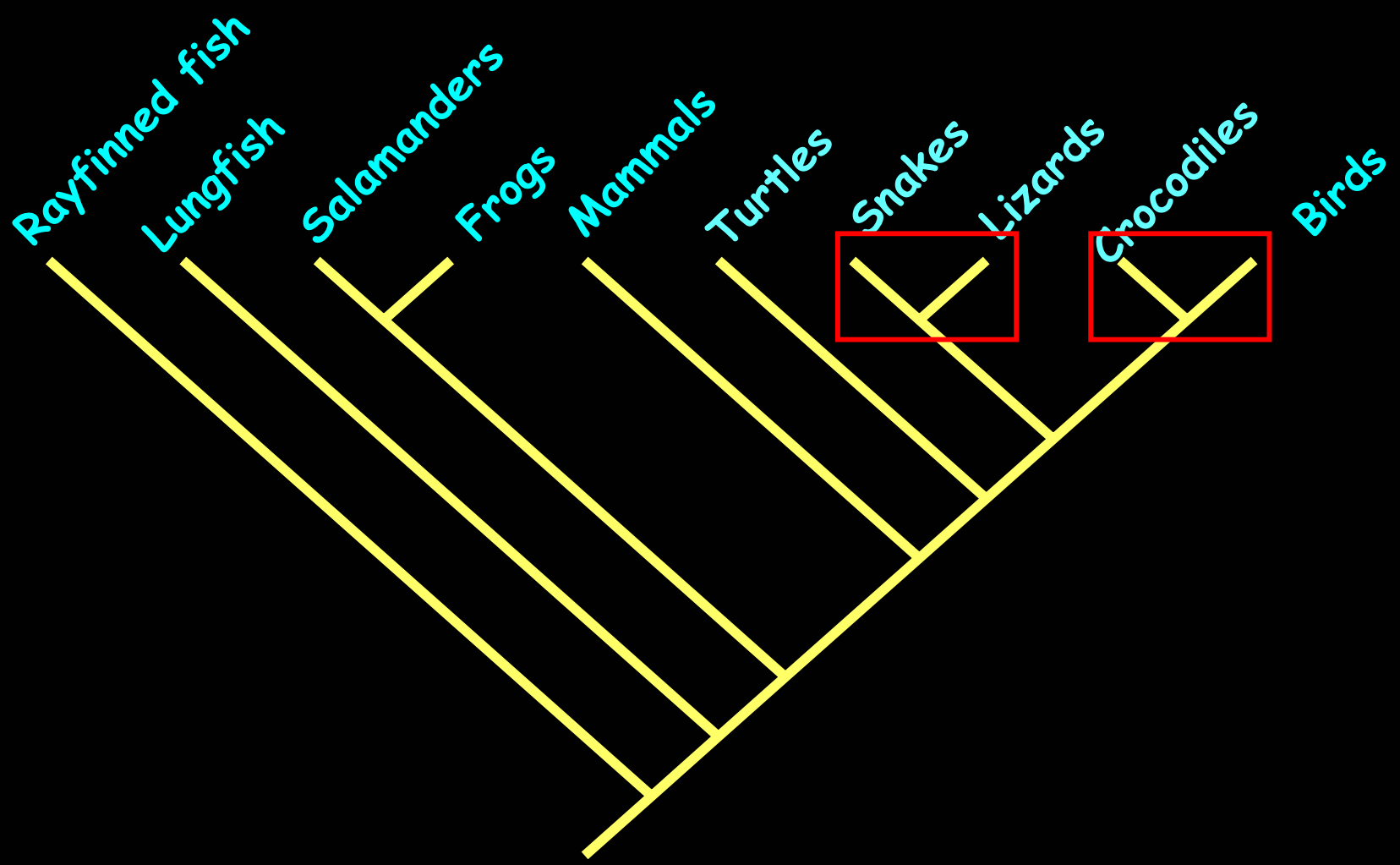
USGS
science for a changing world


Lampreys
Sharks
Teleosts
Lungfish
Turtles
Mammals
Birds
Crocodiles
Snakes
Lizards
Salamanders
Frogs
A
Lampreys
Sharks
Teleosts
Lungfish
Turtles
Mammals
Birds
Crocodiles
Snakes
Lizards
Salamanders
Frogs
B
Lampreys
Sharks
Teleosts
Lungfish
Turtles
Mammals
Birds
Crocodiles
Snakes
Lizards
Salamanders
Frogs
C
Sharks
Lampreys
Teleosts
Lungfish
Frogs
Salamanders
Turtles
Mammals
Lizards
Snakes
Crocodiles
Birds
D

Why phylogenetics?

- Enables the elucidation of natural (=monophyletic) groups and natural classification

USGS
science for a changing world

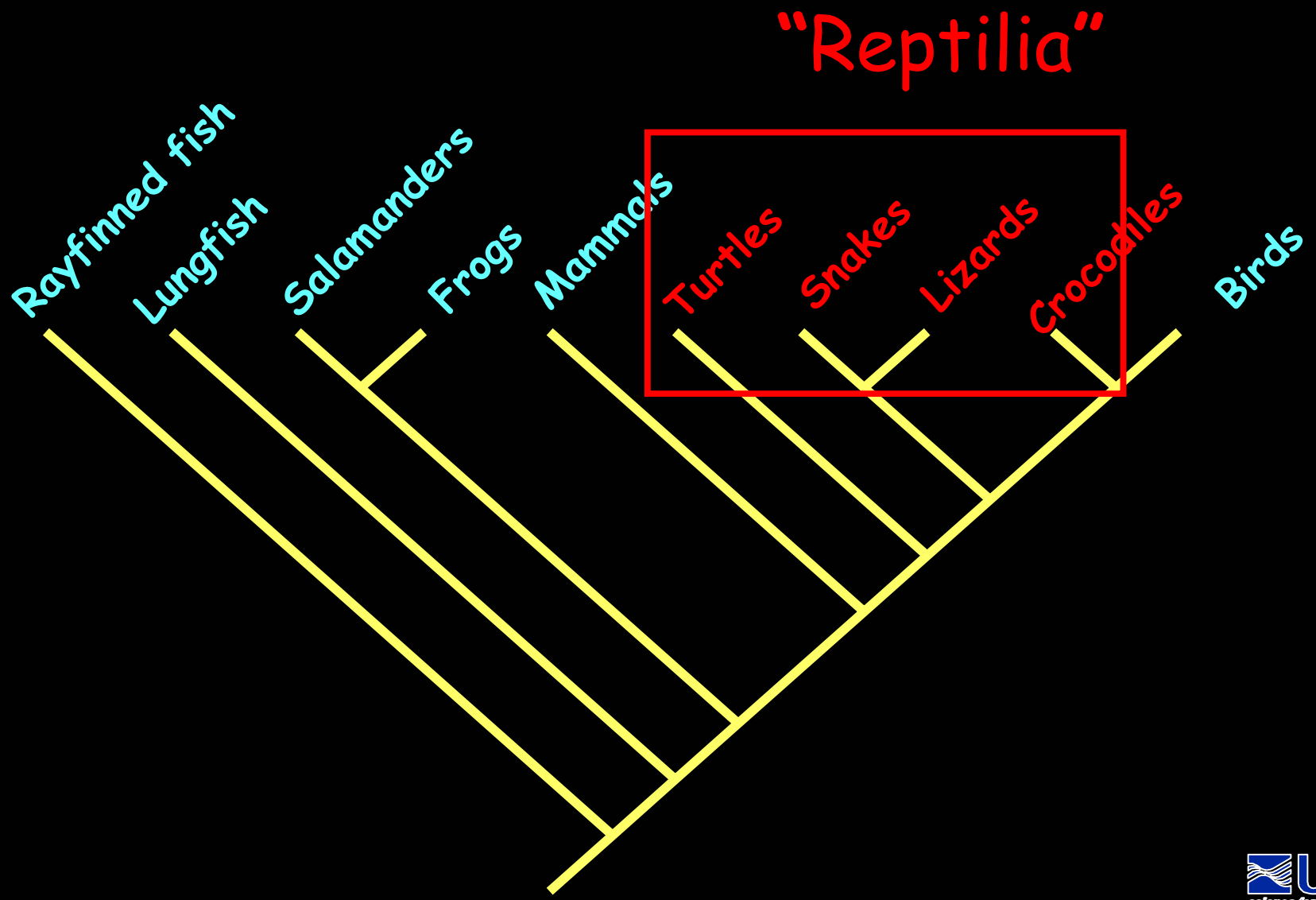


USGS
science for a changing world

Why phylogenetics?

- Enables the elucidation of natural groups and natural classification
- Enables the most efficient possible information retrieval system

USGS
science for a changing world



USGS
science for a changing world

Why phylogenetics?

- Enables the elucidation of natural groups and natural classification
- Enables the efficient possible information retrieval system
- Enables inference of the sequence and number of evolutionary events and the testing of historical evolutionary hypotheses

USGS
science for a changing world

So what?

- Enables the testing of historical hypothesis
  - Adaptational hypotheses
  - Biogeographic hypotheses
  - Behavioral hypotheses

USGS
science for a changing world

Why phylogenetics?

- Enables the elucidation of natural groups and natural classification
- Enables the efficient possible information retrieval system
- Enables inference of the sequence and number of evolutionary events and the testing of historical evolutionary hypotheses
- Enables understanding and prioritization of conservation priorities

USGS
science for a changing world

Alternative classifications

- "Evolutionary" taxonomists would accept a demonstrably paraphyletic group (e.g. Reptilia)

- "Cladists" accept only monophyletic groups in classification

USGS
science for a changing world

History

- Early debates surrounded "schools" of systematics: Evolutionary taxonomists, "numerical taxonomists" (who embraced phenetic methods), and phylogenetic systematists (cladists)
- Primary contributions of phylogenetic systematics: Monophyly and polarity (apomorphy). More broadly: An efficient information retrieval system with the potential to reflect evolutionary history in classification

USGS
science for a changing world

Key Concepts

- Monophyly:
  Tree-based term

- Polarity (apomorphy):
  Character-based term

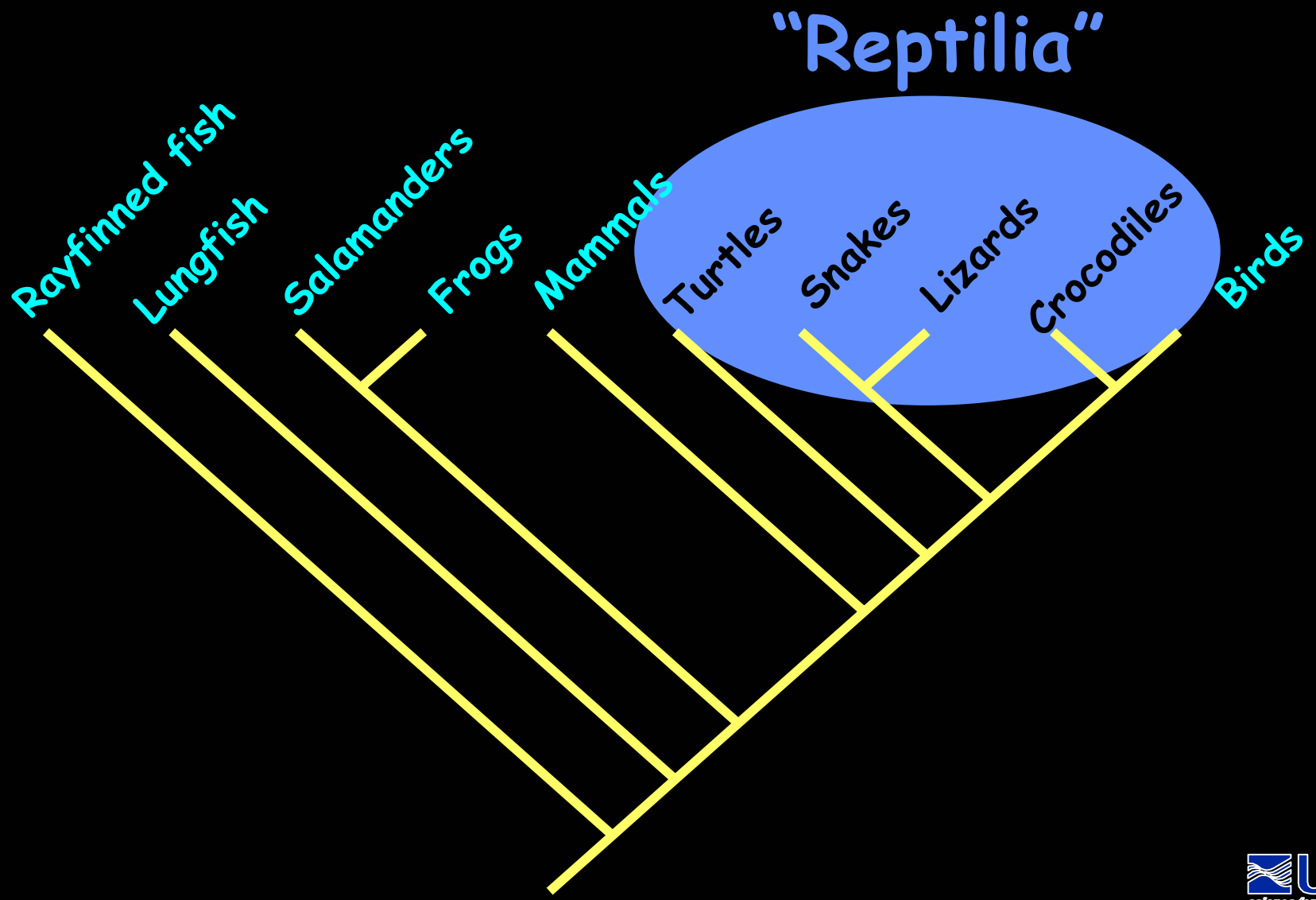
USGS
science for a changing world

Monophyletic Group

- A group that includes a common ancestor and all of its descendants
- A group with unique and unreversed group membership characters

USGS
science for a changing world

Monophyly and polarity



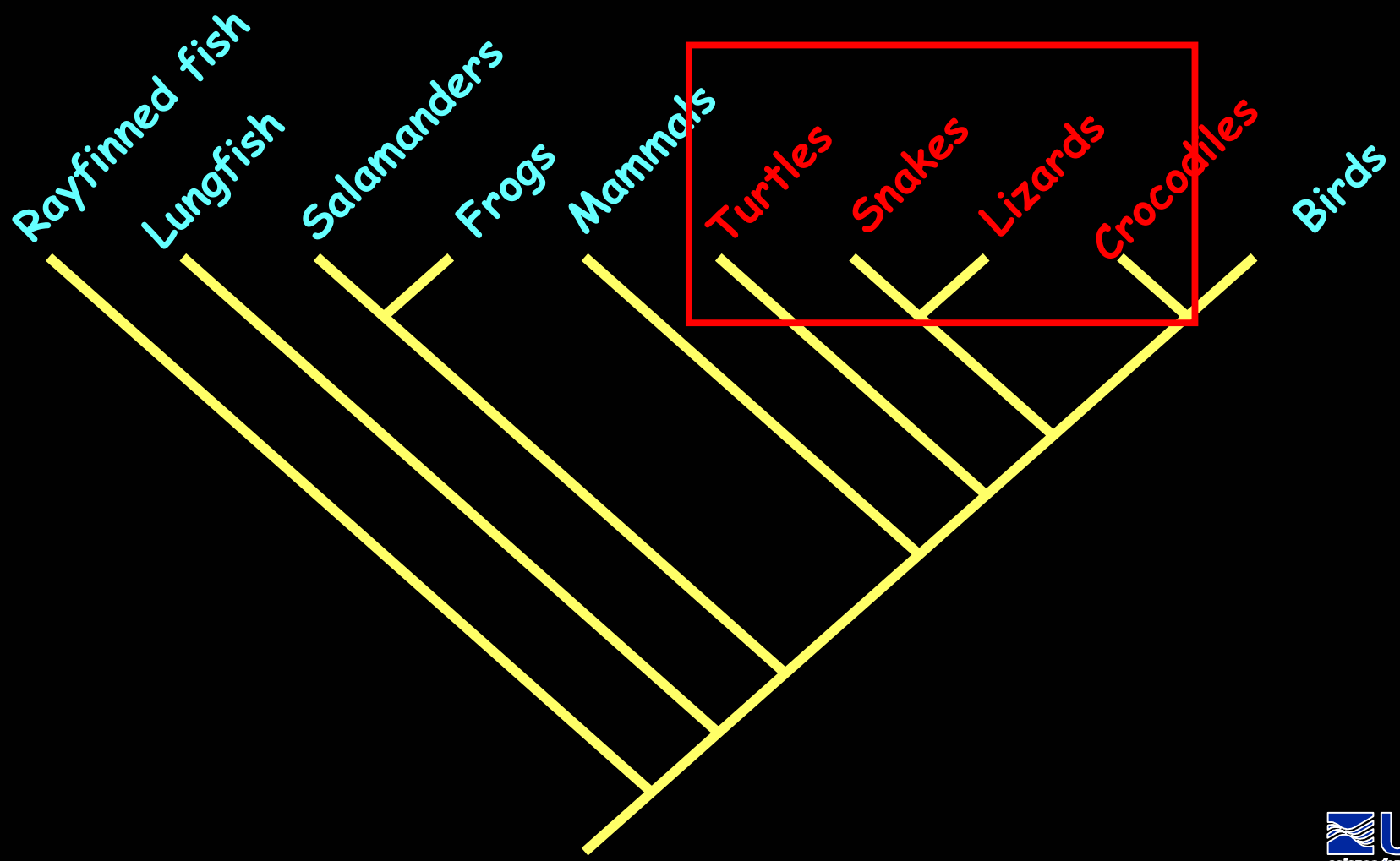
USGS
science for a changing world

Paraphyletic Group

- A group that includes a common ancestor and some but not all of its descendants
- A group with unique but reversed group membership characters

USGS
science for a changing world

Paraphyletic Group: "Reptilia"



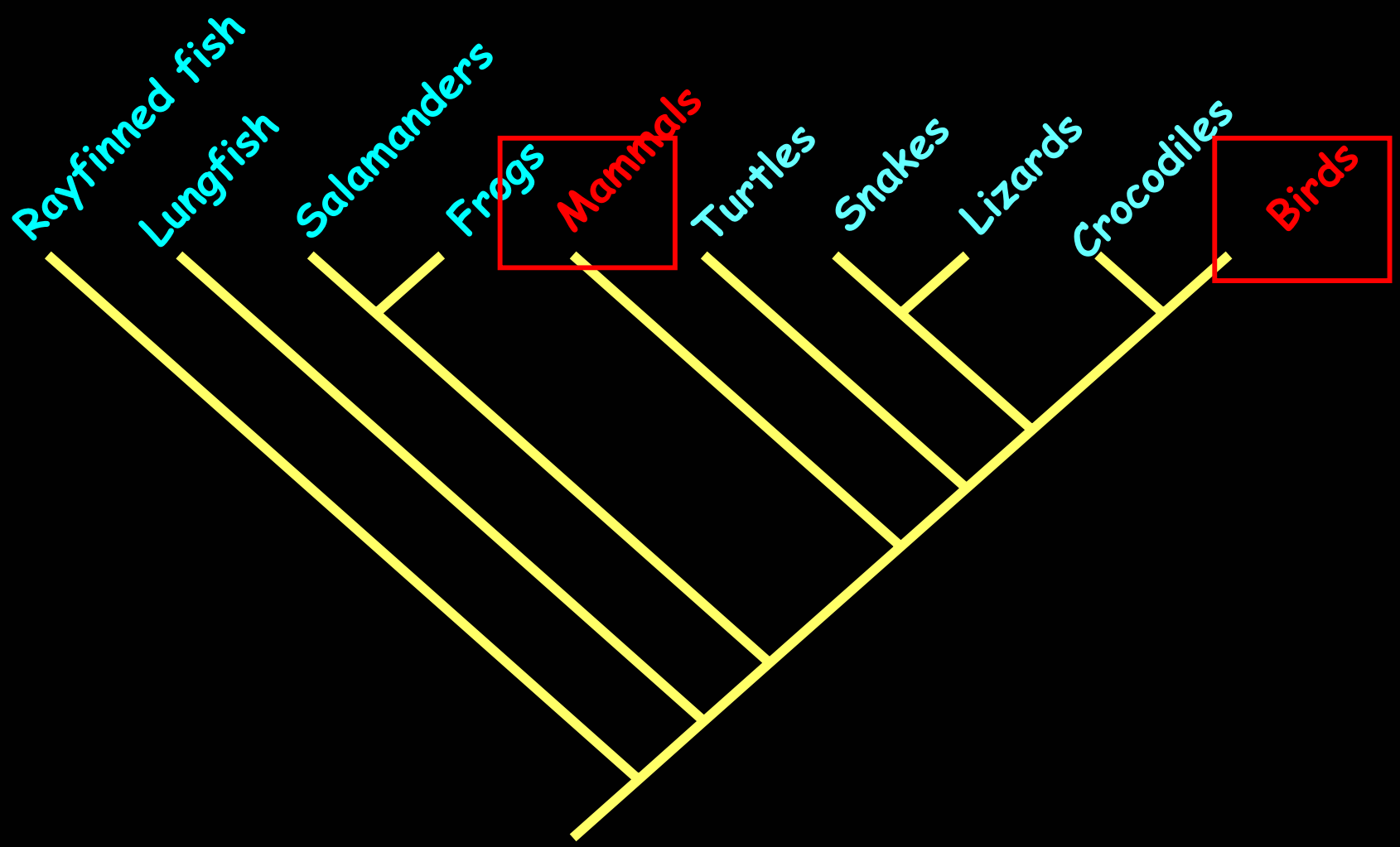
USGS
science for a changing world

Polyphyletic Group

- A group in which the most recent common ancestor is assigned to some other group
- A group whose membership characteristics are not uniquely derived

USGS
science for a changing world

Polyphyletic Group: Haemothermia

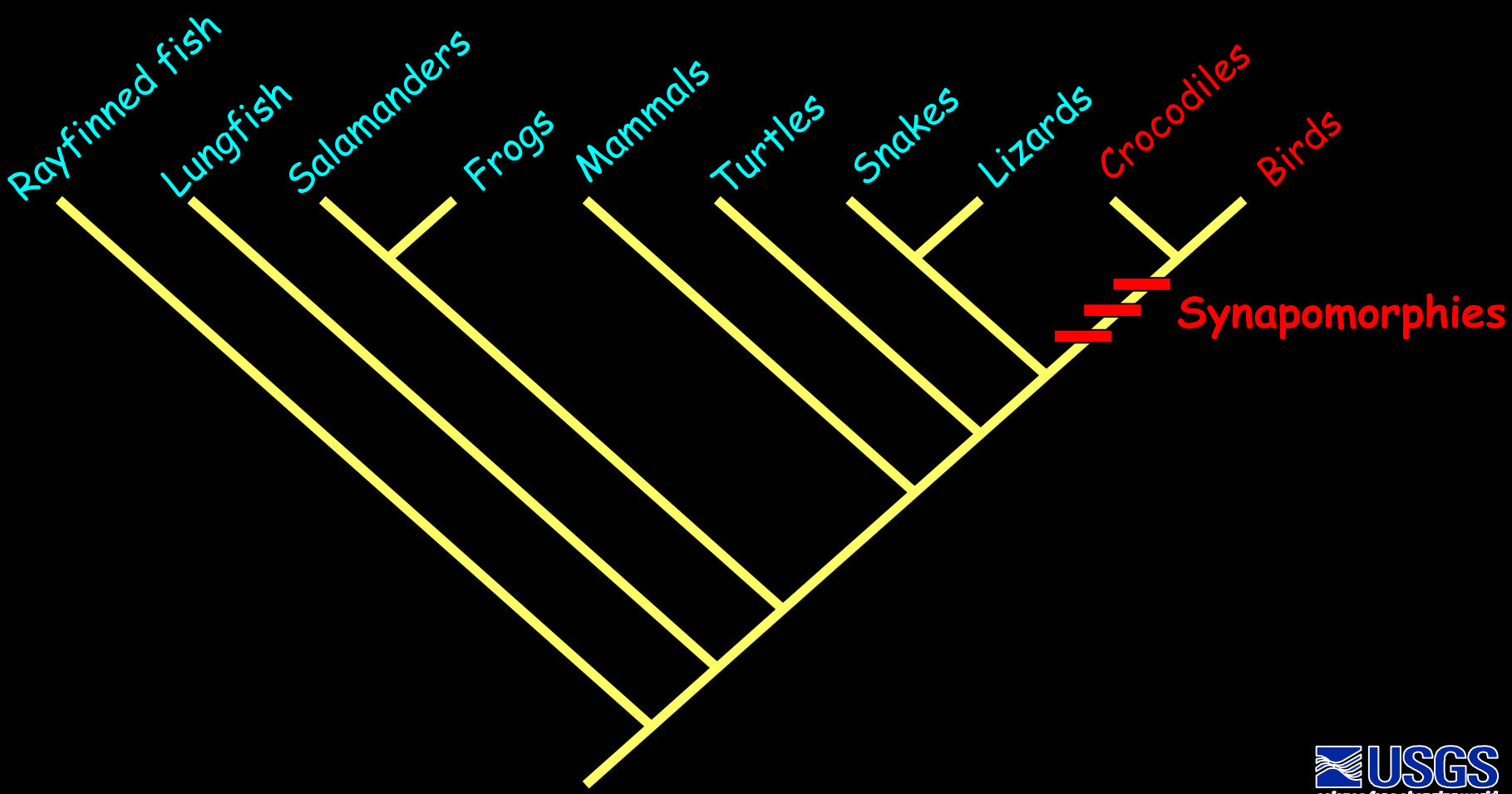


USGS
science for a changing world

Polarity-related terms

- Apomorphy: Derived character
  - Synapomorphy: Shared derived character
  - Autapomorphy: Uniquely derived character (no information) [example – tetrapod limbs]

USGS
science for a changing world



USGS
science for a changing world

Polarity-related terms

- Apomorphy: Derived character
  - Synapomorphy: Shared derived character
  - Autapomorphy: Uniquely derived character (no information)
- Plesiomorphy: "Primitive" character
  - Symplesiomorphy: Shared primitive character [example – lizard limbs]

USGS science for a changing world

Polarity-related terms

- Apomorphy: Derived character
  - Synapomorphy: Shared derived character
  - Autapomorphy: Uniquely derived character (no information) [example – tetrapod limbs]
- Plesiomorphy: "Primitive" character
  - Symplesiomorphy: Shared primitive character [example – lizard legs]
- Homoplasy: Parallelism, reversal, convergence [example – vision]

USGS
science for a changing world

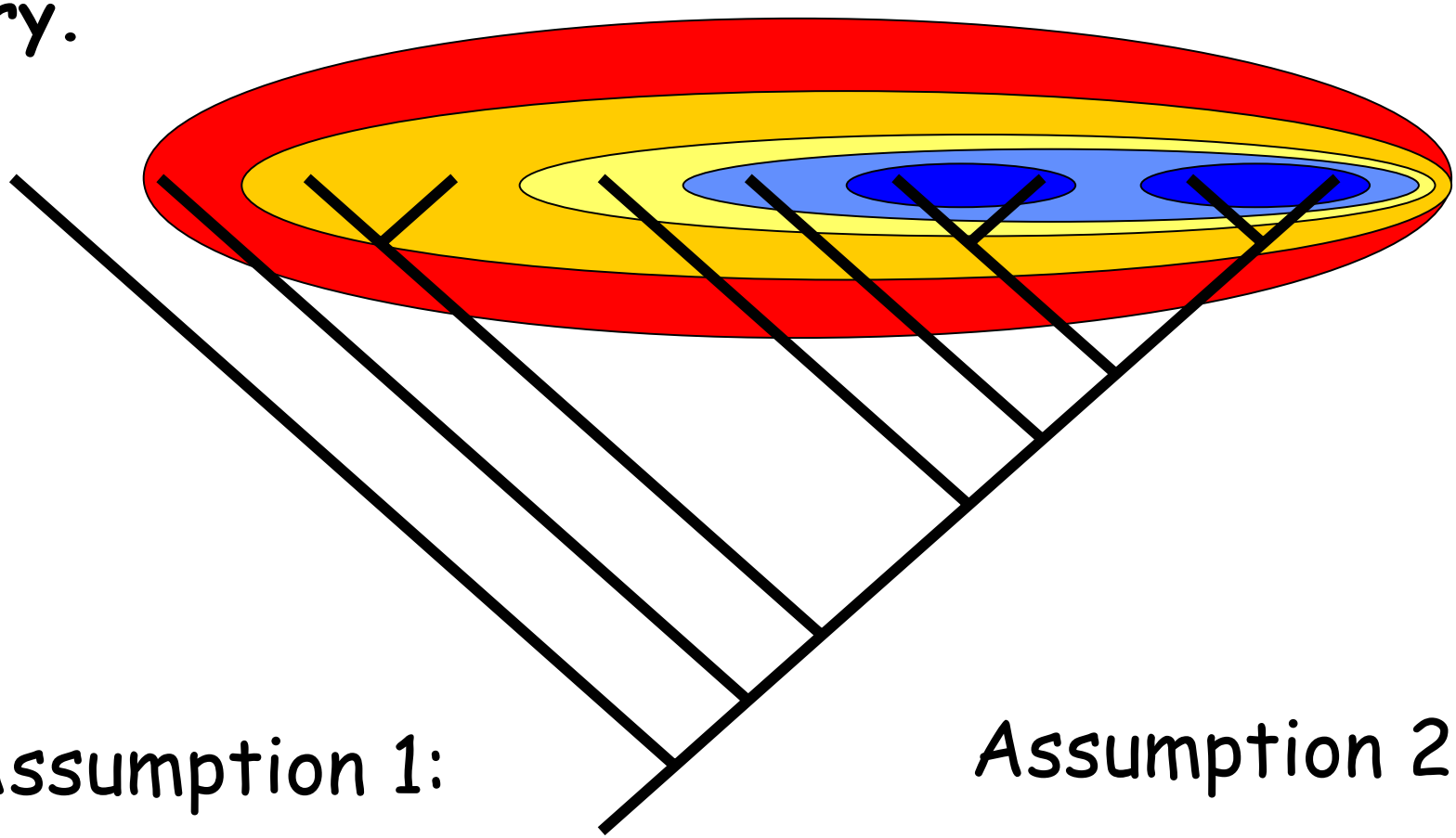
A cladogram depicts relative recency of common ancestry.

No more

No less

USGS
science for a changing world

A cladogram depicts relative recency of common ancestry.



Assumption 1:

Descent with modification

Assumption 2:

Cladogenesis

USGS
science for a changing world

History

- Ultimately, as most agree that phylogenetics are important and interesting, the debates have shifted to phylogenetic inference methods

USGS
science for a changing world

History

- Ultimately, as most agree that phylogenetics are important and interesting, the debates have shifted to phylogenetic inference methods

- To a certain degree, these debates have proceeded at cross purposes

USGS
science for a changing world

The best one can hope to do is be aware of the issues, understand the methods in order to evaluate critically a given data set's relevance to a conservation question

USGS
science for a changing world

Phylogenetic Inference/Reconstruction Methods

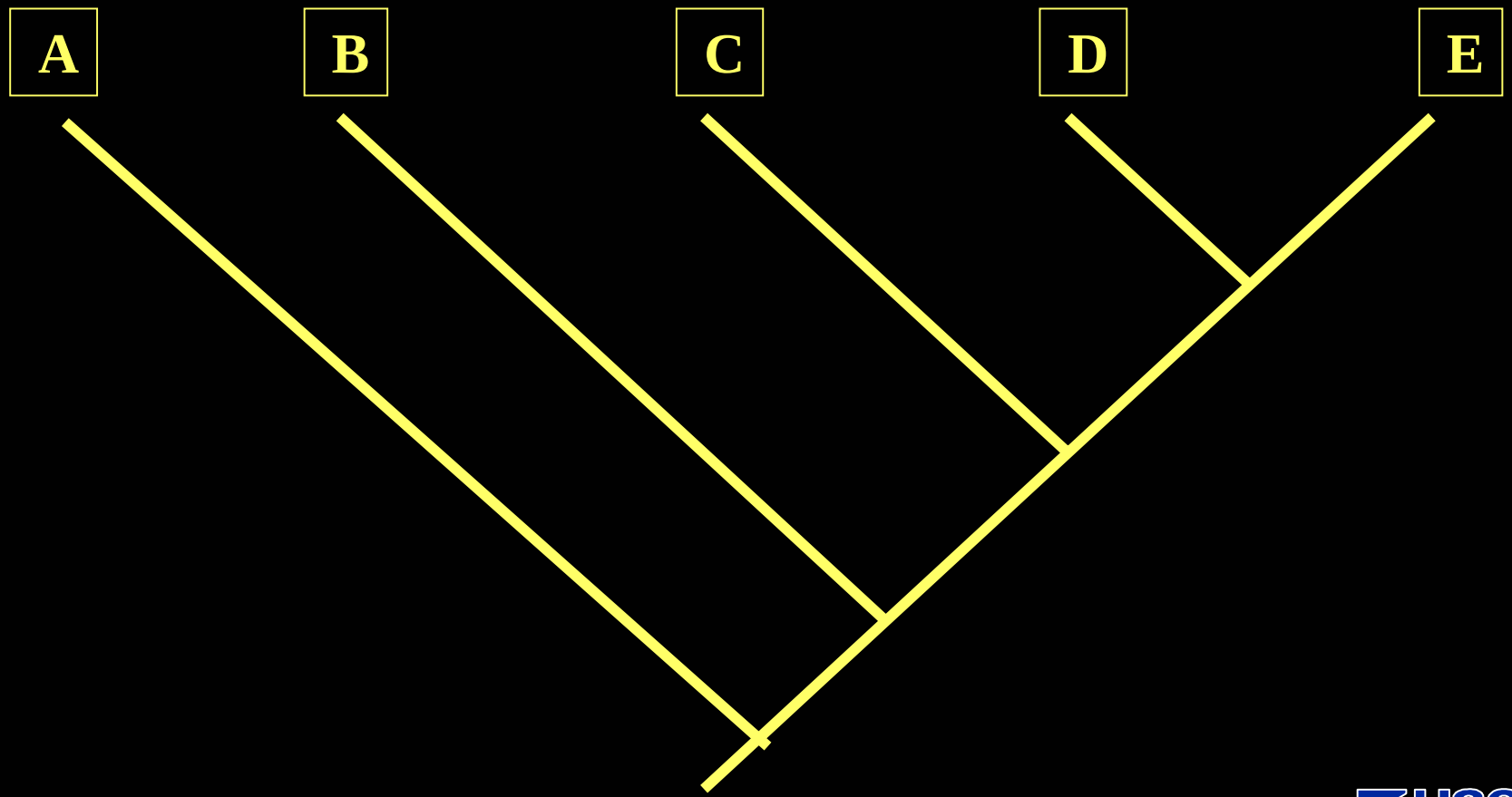
- How do you go from this…

Taxa	Characters					
A	0	0	0	0	0	1
B	0	0	0	0	1	1
C	0	0	0	1	1	1
D	0	0	1	1	1	1
E	0	1	1	1	1	1

USGS
science for a changing world

Phylogenetic Inference/Reconstruction Methods

- …to this



USGS
science for a changing world

Phylogenetic Inference/Reconstruction Methods

Taxa	Characters					
A	0	0	0	0	0	1
B	0	0	0	0	1	1
C	0	0	0	1	1	1
D	0	0	1	1	1	1
E	0	1	1	1	1	1

USGS
science for a changing world

Phenetic Techniques

- Clustering methods based on overall similarity

USGS
science for a changing world

Phenetic Analyses

- WPGMA

USGS
science for a changing world

	perch	coelocanth	salamander	frog	turtle	human	gecko	snake	alligator	budgy
perch		85	62	62	69	54	54	62	54	23
coelocanth			77	77	85	69	69	77	69	38
salamander				100	77	77	54	46	62	46
frog					77	77	54	46	62	46
turtle						85	85	76	85	54
human							69	62	69	69
gecko								92	85	54
snake									76	46
alligator										69
budgy										

USGS
science for a changing world

WPGMA

The similarity between perch and (salamander, frog) is
(perch:salamander+perch:frog+salamander:frog)/3, or
= (62+62+100)/3
= 75
and then, too, the similarity between (salamander, frog) and gecko is then
(salamander:gecko+frog:gecko+ salamander:frog)/3
= (54+54+100)/3
= 69

USGS
science for a changing world

Phenetic Techniques

- DNA-DNA hybridization

USGS
science for a changing world

Distance Analyses

- Neighbor-joining

USGS
science for a changing world

Issues with Phenetics

- Metricity, negative branch lengths, triangle inequality

USGS
science for a changing world

Optimality Criteria

USGS
science for a changing world

Phylogenetic Inference/Reconstruction Methods

- Parsimony: Justified in part by Hennig's so-called Auxiliary Principal, namely that we not hypothesize homoplasies (convergences, parallelisms, reversals) beyond what are required by the data.

The idea derives from Occam's Razor, ie that the explanation requiring the fewest assumptions (in this case the **shortest tree**) is the most robust, the best corroborated, etc.

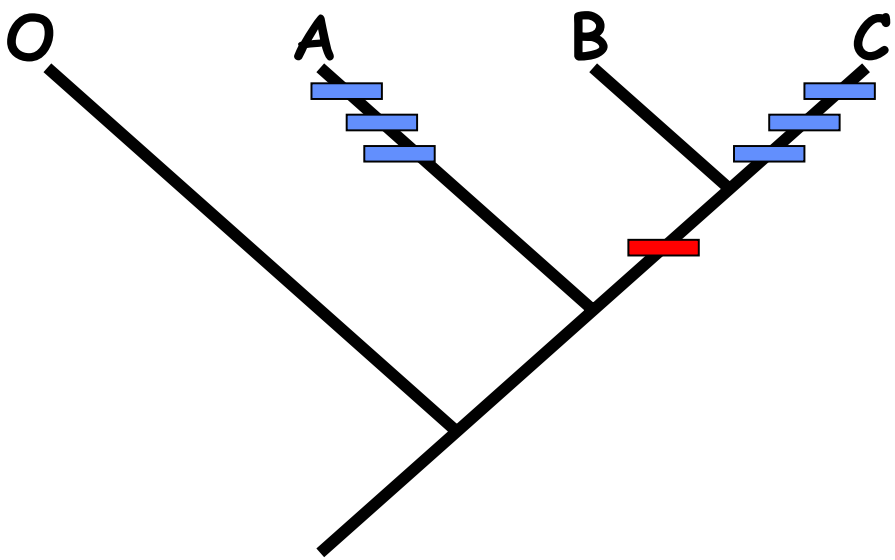
USGS
science for a changing world

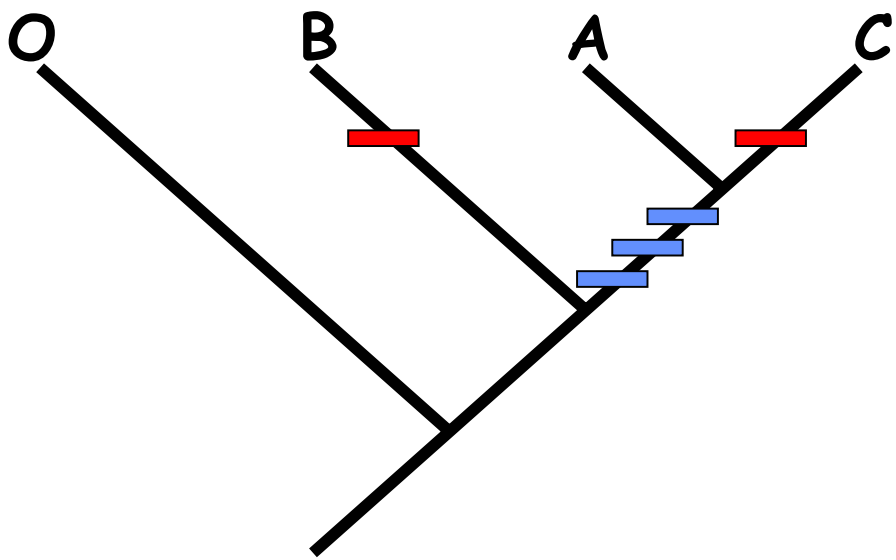
Phylogenetic Inference/Reconstruction Methods

- Parsimony: Justified in part by Hennig's Auxiliary Principal, namely that we not hypothesize homoplasies (convergences, parallelisms, reversals) beyond what are required by the data.
  - Most generalizable; accommodates any kind of phylogenetic data (morphological, molecular, behavioral, etc.) with equal aplomb

USGS
science for a changing world

Parsimony and character support

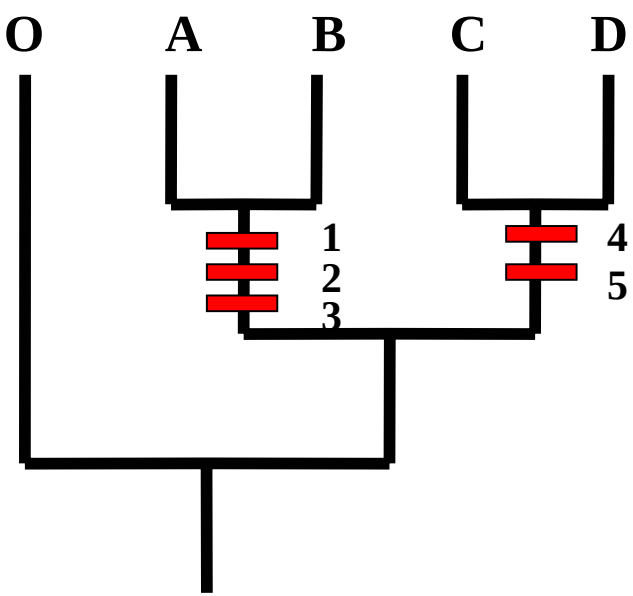




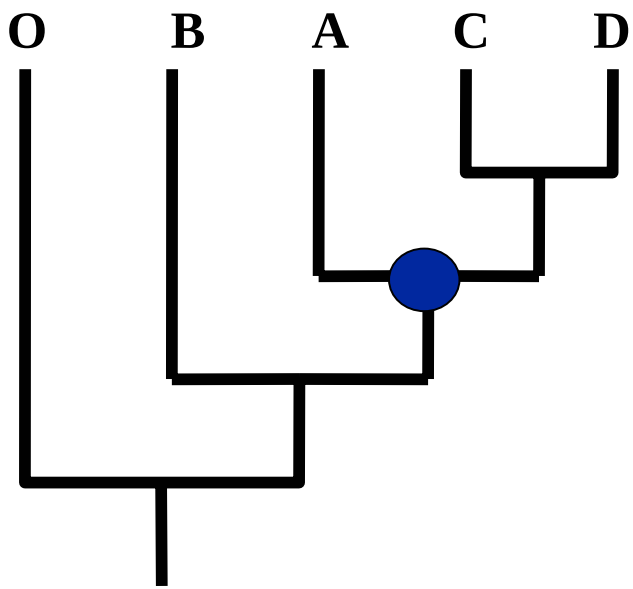
Recall:
Parsing Homoplasy from Synapomorphy`

USGS
science for a changing world

Taxon	12345……..
O	gggaaaaaaaaaaggggg
A	aaataaaaaaaaaagaaa
B	aaatgggggggggagaaa
C	ttttaaaaaaaaaagaatt
D	ctctaaaaaaaaaagataa



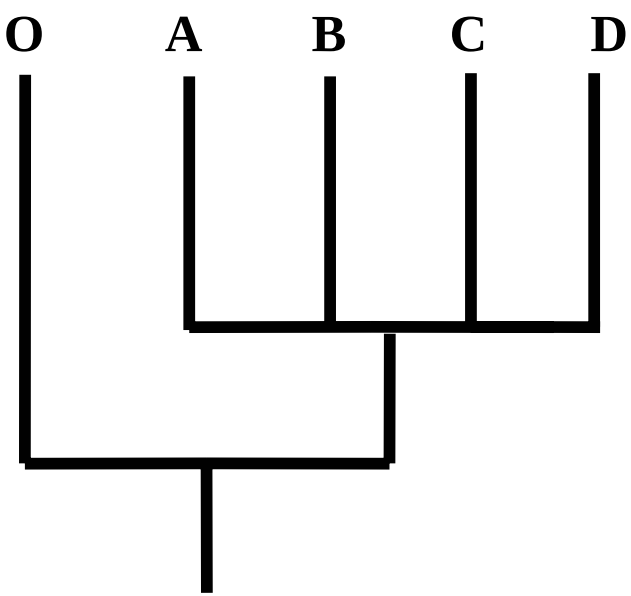
Parsimony



WPGMA

USGS

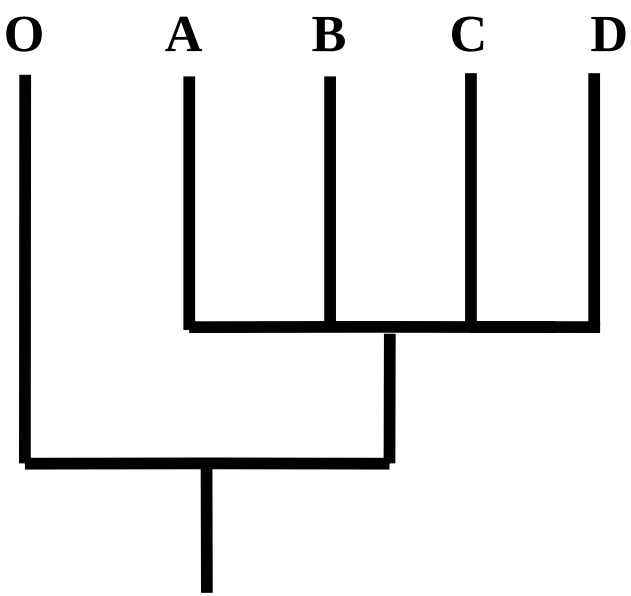
Taxon	1 2 3 4 5 6 7 8 9 10
O	AAAAAAAAAA
A	AAAAAAAAAA
B	AAAAAAAAAA
C	AAAAAAAAAA
D	AAAAAAAAAA



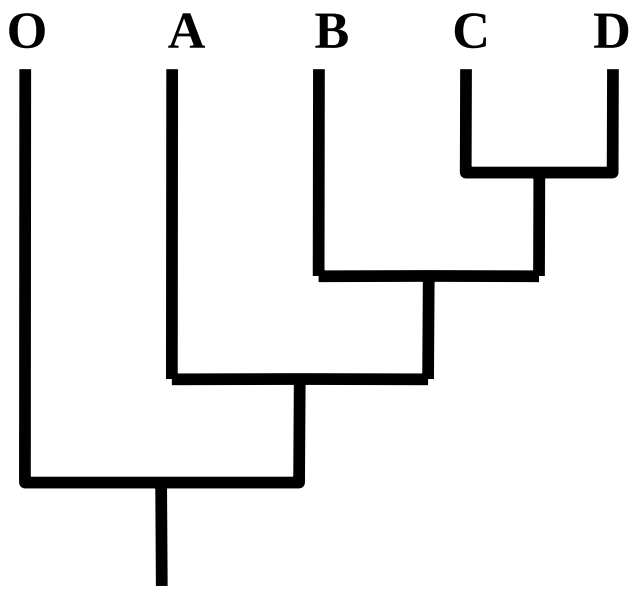
Parsimony

USGS
science for a changing world

Taxon	1 2 3 4 5 6 7 8 9 10
O	AAAAAAAAAA
A	AAAAAAAAAA
B	AAAAAAAAAA
C	AAAAAAAAAA
D	AAAAAAAAAA



Parsimony



Neighbor Joining

USGS
science for a changing world

Rooting

- The primary result of a phylogenetic analysis is a network that does not in and of itself imply exclusive relationship because it does not make polarity statements a priory
- This requires rooting the tree through the explicit inclusion of a suitable outgroup

USGS
science for a changing world

USGS
science for a changing world

Branch Support

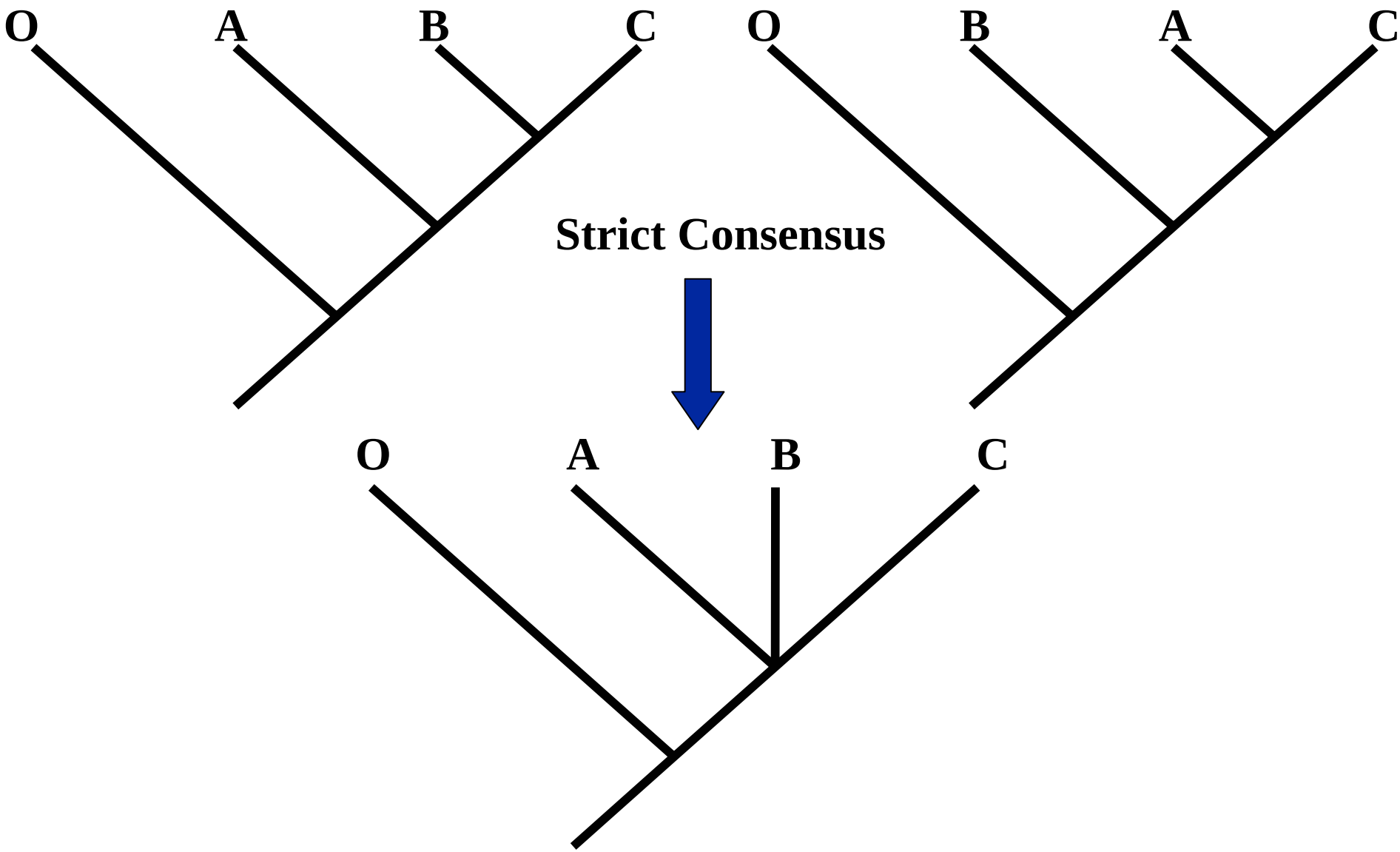
- Straightforward character mapping
- Bootstrap values
- Bremer values

USGS
science for a changing world

Dealing with data, results

- The most parsimonious tree is the tree that requires the fewest steps. Multiple equally short trees may obtain. This reflects indecisive data.

USGS
science for a changing world


O
A
B
C
O
B
A
C
Strict Consensus
O
A
B
C

Dealing with data, results

- The most parsimonious tree is the tree that requires the fewest steps. Multiple equally short trees may obtain. This reflects indecisive data.

- Interpreting lack of resolution

USGS
science for a changing world

Morphology and Molecules

- Qualitatively different

USGS
science for a changing world

Morphology and Molecules

- Qualitatively different
- Quantitatively the same?

USGS
science for a changing world

Patterns of Substitution

- Silent versus replacement substitutions
- 1^{st}, 2^{nd}, and 3^{rd} positions
- Transitions (purine >>>purine or pyrimidine>>>pyrimidine) and transversions (purine<<<>>>pyrimidine)

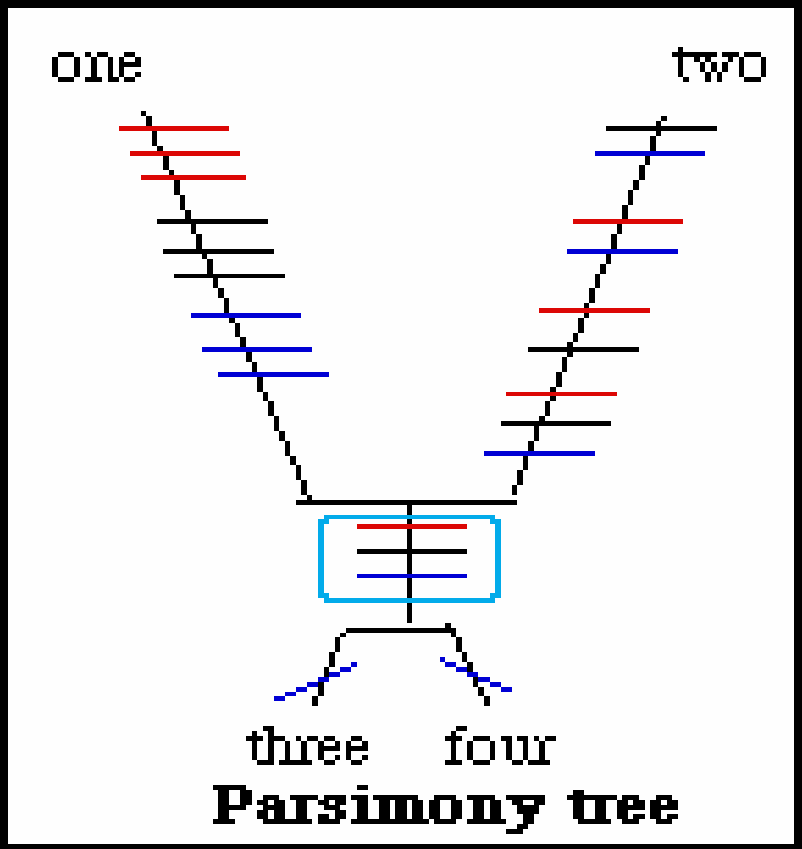
USGS
science for a changing world

Phylogenetic Inference/Reconstruction Methods

- Parsimony: Justified in part by Hennig's so-called Auxiliary Principal, namely that we not hypothesize homoplasies (convergences, parallelisms, reversals) beyond what are required by the data
- Maximum Likelihood: Model-based method deriving in part from justification that parsimony may not always perform with statistical consistency (the long branch attraction epi-phenomenon)

USGS
science for a changing world

A AAAACCCC TTTT
CAC TAC TAC TAC T
AAAAAAAAAAAAAAAA
CAAAAAAAAAAAAAAA



USGS
science for a changing world

ML Seeks to Incorporate

- Recognized molecular evolutionary principles
- Frequencies of change and rates of evolution implied by those principles

So as to "weight" characters according to their inferred importance

USGS
science for a changing world

Models

- Parameterization
- Complication

USGS
science for a changing world

Phylogenetic Inference/Reconstruction Methods

- Bayesian Analysis:

  The Monte Hall Problem (credited to W. Wheeler) and the use of prior probabilities
- Asserted to be generalizable to all forms of data with similar model-able features as strict maximum likelihood

USGS
science for a changing world

Debates

- Probabilism versus parsimony

USGS
science for a changing world

Molecular "versus" Morphological Data

- Historical and current debate with strong advocates on both sides

USGS
science for a changing world

Molecular "versus" Morphological Data

Molecular proponents:

- Morphology is subject to convergence, has contributed nothing to systematics or is otherwise irrelevant (Graur, Hedges)
- Molecular data constitute an "independent" test (!)
- Concerns over data "swamping"

USGS
science for a changing world

Molecular "versus" Morphological Data

Morphological proponents:

- Molecular data do not feed the process of systematics adequately (e.g. with reference to character circumscription or, commonly, type species)
- Commonly practiced molecular analyses do not enable an understanding an understanding of synapomorphy or character support

USGS
science for a changing world

Morphology “versus” molecules redux

- Molecular data are not superior to morphological data

USGS
science for a changing world

Morphology “versus” molecules redux

- Molecular data are not superior to morphological data
- Morphological data are not superior to molecular data

USGS
science for a changing world

Morphology “versus” molecules redux

- Molecular data are not superior to morphological data
- Morphological data are not superior to molecular data
- No class of data is inherently superior to any other class of data; rather they complement one another if analyzed properly.

USGS
science for a changing world

Morphology “versus” molecules redux

- Molecular data are not superior to morphological data
- Morphological data are not superior to molecular data
- No class of data is inherently superior to any other class of data; rather they complement one another if analyzed properly.
- Molecular data provide a fantastically rich source of phylogenetic information

USGS
science for a changing world

Morphology "versus" molecules redux

- Molecular data are not superior to morphological data
- Morphological data are not superior to molecular data
- No class of data is inherently superior to any other class of data; rather they complement one another if analyzed properly.
- Molecular data provide a fantastically rich source of phylogenetic information
- Morphological (and behavioral and ecological) data represent the foundation on which the basic description of the natural world rests

USGS
science for a changing world

Combinability: Options

- Simultaneous analysis of all data sets

USGS
science for a changing world

Combinability: Options

- Simultaneous analysis of all data sets
- *A* priori partitioning of data sets followed by separate analyses and post hoc decisions as to how to deal with discrepancies

USGS
science for a changing world

Combinability: Options

- Simultaneous analysis of all data sets
- *A* priori partitioning of data sets followed by separate analyses and post hoc decisions as to how to deal with discrepancies
- Testing for heterogeneity or incongruity

USGS
science for a changing world

- A chainsaw, like any good tool can be applied towards many useful purposes
- Likewise genetic data may be used to answer a variety of questions that require different strategies of hypothesis testing, data gathering, and analysis
- Genetics is a meaningless term in and of itself unless contextualized

USGS
science for a changing world

Background

- "Genetics" used variously in conservation
- Distinction between population viability-related concerns and inferring historical relationships among management units

USGS
science for a changing world

Salient Points

- Methodological issues in phylogenetic analysis and delineation of conservation "units"
- Correspondence (or lack thereof) of delineated management units and taxonomic epithets
- In the context of delineating units (intra-specific and otherwise), genetic data offer a promising set of characters, but they are not the only characters, or necessarily the most relevant

USGS
science for a changing world

- A chainsaw can also be applied irresponsibly

USGS
science for a changing world

Molecular Ecology (2004) 13, 1759–1769 doi: 10.1111/j.1365-294X.2004.02181.x

Genetic evaluation of a proposed introduction: the case of the greater prairie chicken and the extinct heath hen

ERIC P. PALKOVACS,* ADAM J. OPPENHEIMER,* EUGENE GLADYSHEV,† JOHN E. TOEPFER,‡ GEORGE AMATO,§ THOMAS CHASE ¶ and ADALGISA CACCONE*,**

*Department of Ecology and Evolutionary Biology, †School of Forestry and Environmental Studies and **Molecular Systematics and Conservation Genetics Laboratory, Yale Institute for Biospheric Studies, Yale University, 165 Prospect Street, New Haven, CT 06520, USA, ‡Society of Tympanuchus Cupido Pinnatus, Plover, WI 54467, USA, §Wildlife Conservation Society, Bronx Park, NY 10460, USA, ¶The Nature Conservancy, Vineyard Haven, MA 02568, USA

USGS science for a changing world

"Tom Chase hasn't given up…"

Dunlop, 2004: 51

USGS
science for a changing world

Punchlines

- The notion of subspecies can lends itself to misinterpretation (Zink, 2004) in a phylogenetic framework
- Molecular, morphological, ecological, and behavioral data often corroborate one another

USGS
science for a changing world

Phylogeography

- The application of phylogenetic graph theory to population-level biogeographic questions through the medium of mitochondrial DNA

USGS
science for a changing world

Phylogeography

- The uniparental inheritance of mitochondrial DNA means that the relationships among mitochondria are hierarchical, much like the relationships among species and taxonomic groups

USGS
science for a changing world

Phylogeography

- This enables the same kinds of powers cladograms offer the study of species to the study of populations, the idea being that recent evolutionary history can be represented as a nested hierarchy rather than a reticulate network of individual relationships

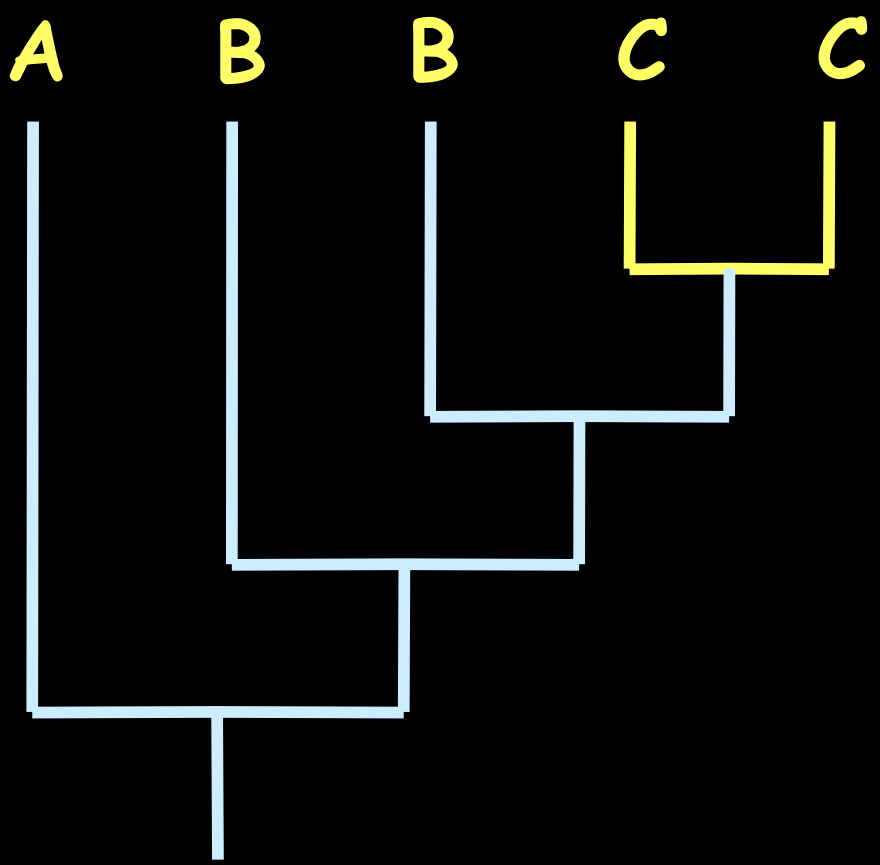
USGS
science for a changing world

Phylogeography

- And so the question becomes whether the relationships among mitochondria necessarily reflect the relationships among populations, species, or even individuals

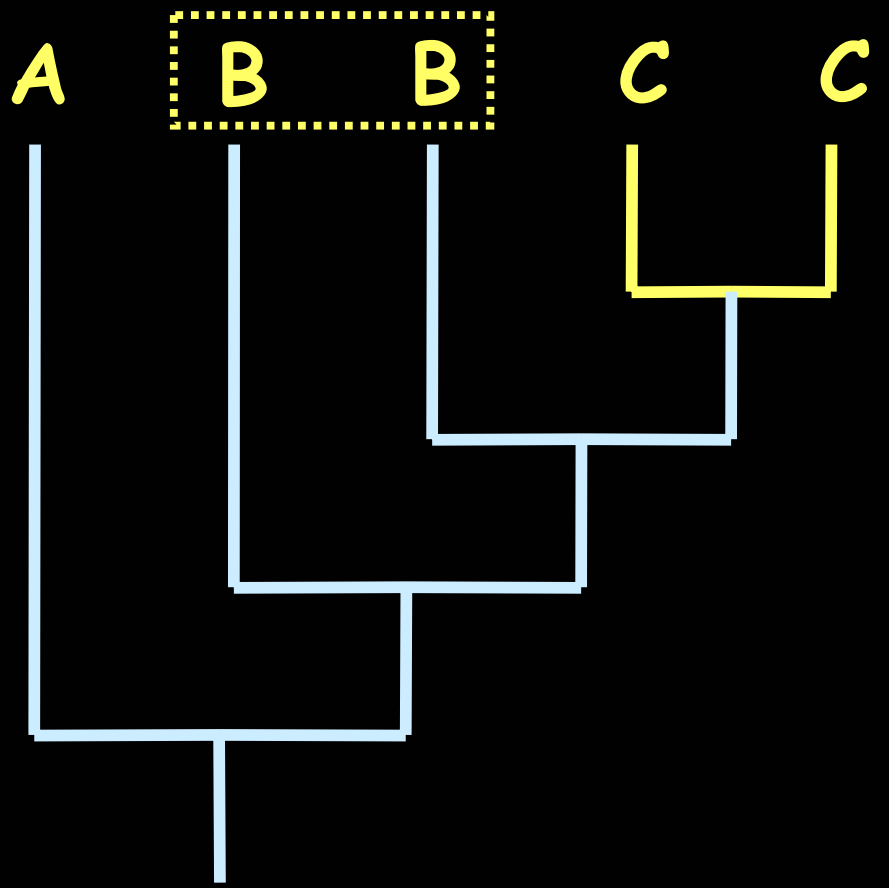
USGS
science for a changing world

Tree-based Conceptions



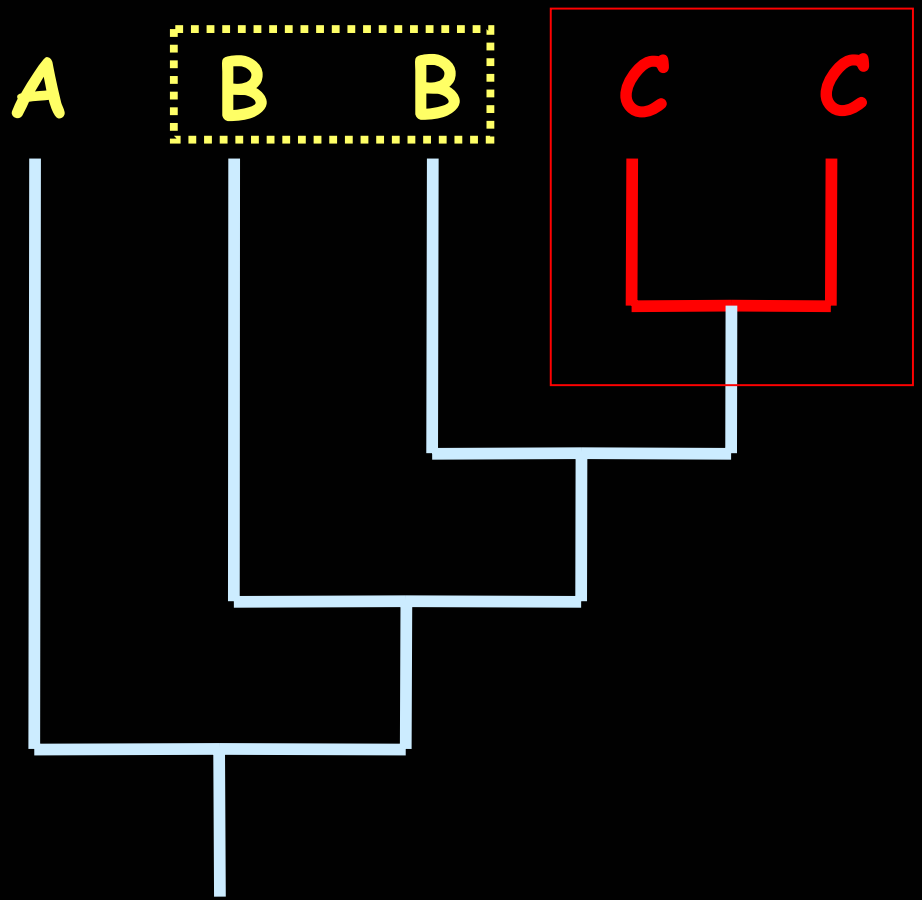
USGS
science for a changing world

Tree-based Conceptions



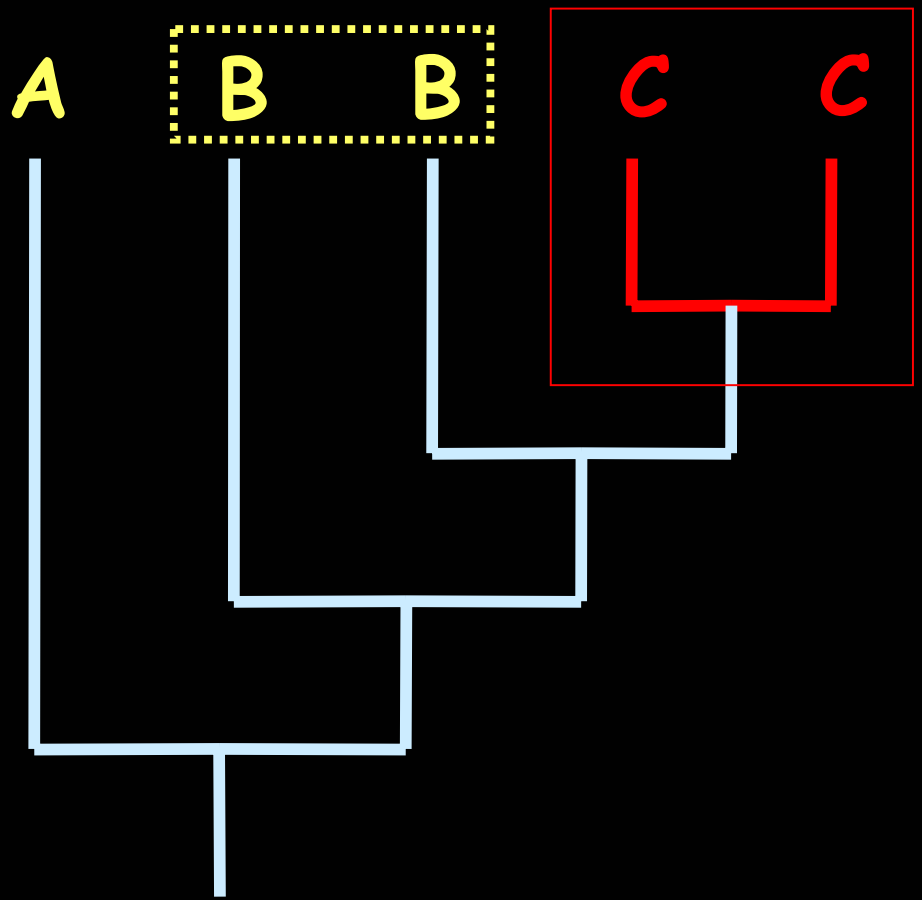
USGS
science for a changing world

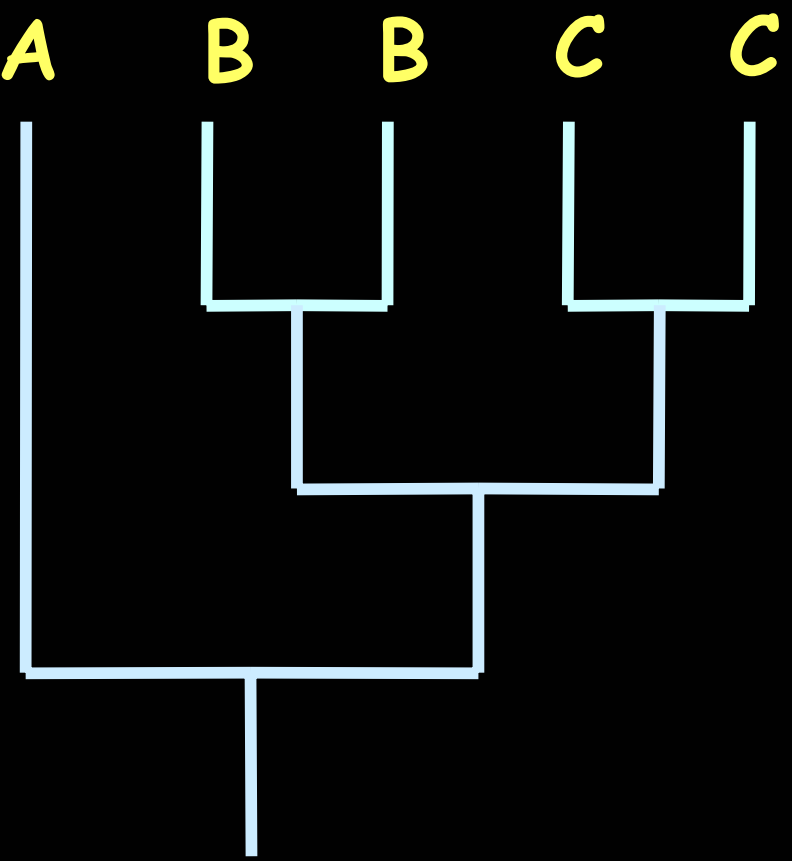
Tree-based Conceptions



USGS
science for a changing world

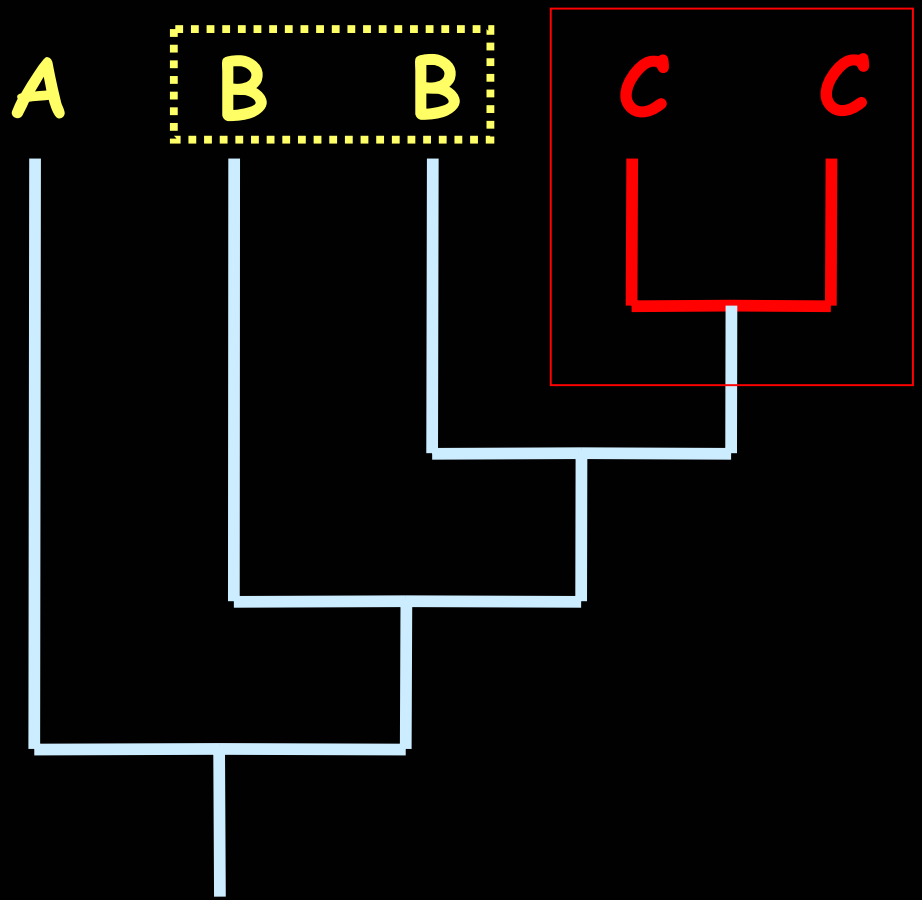
Tree-based Conceptions

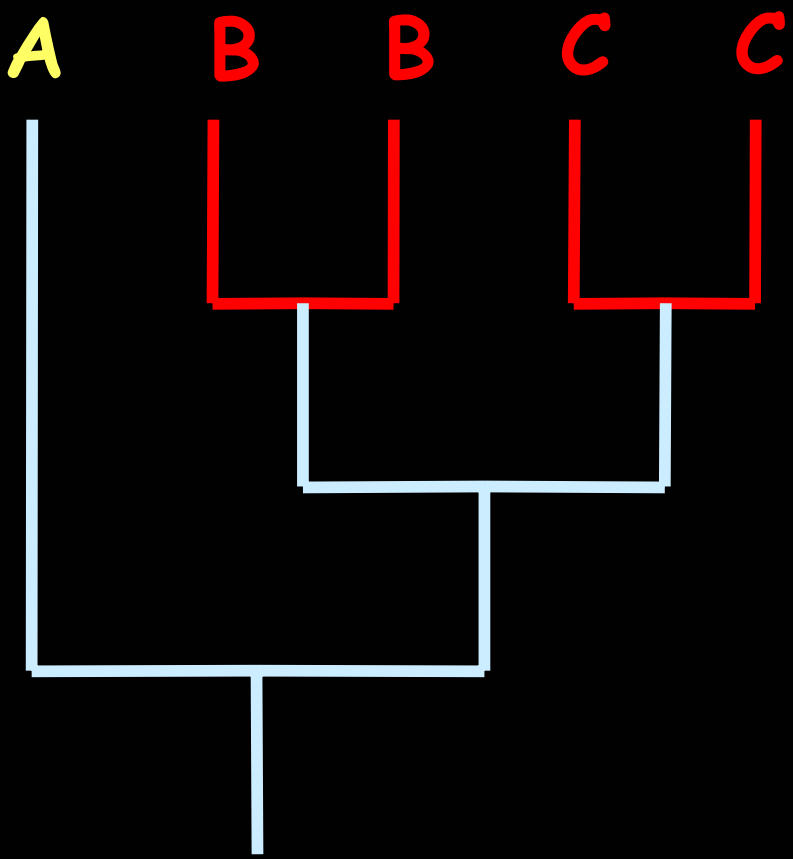




USGS
science for a changing world

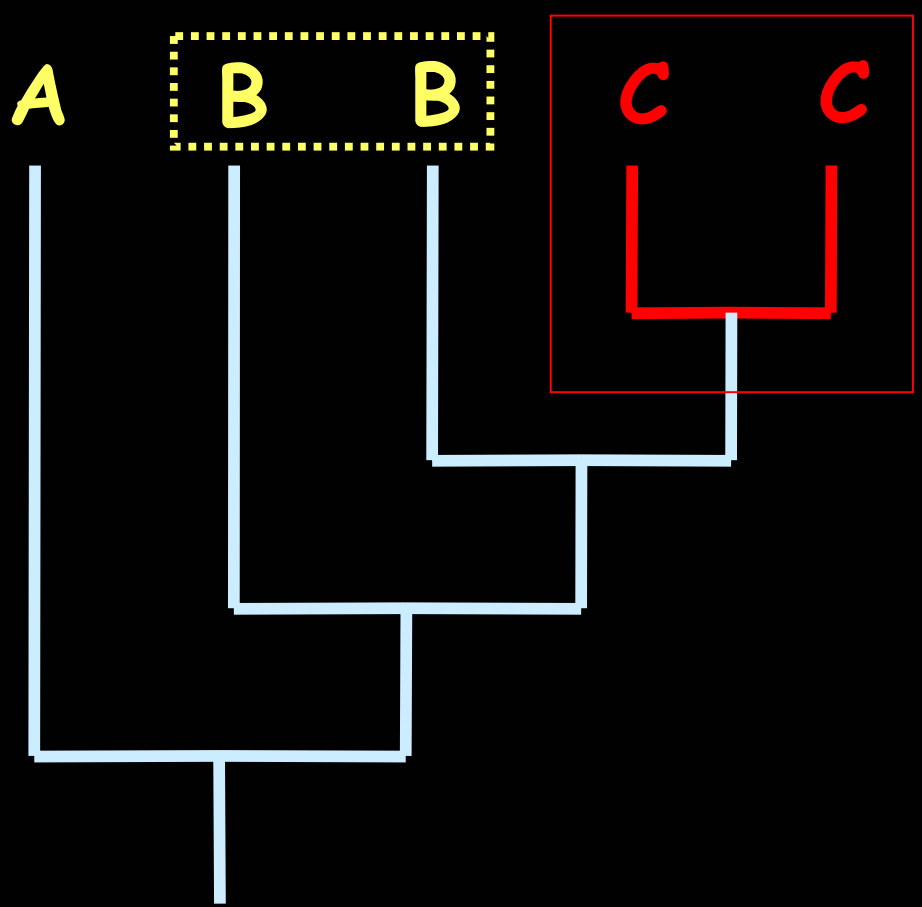
Tree-based Conceptions

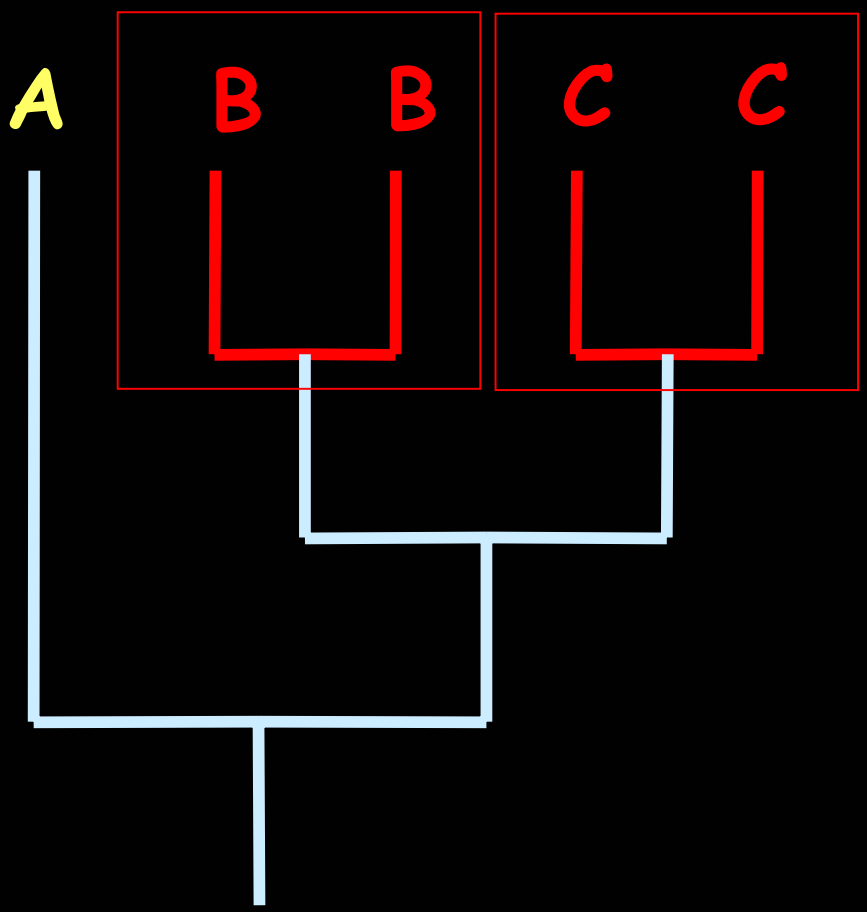




USGS
science for a changing world

Tree-based Conceptions



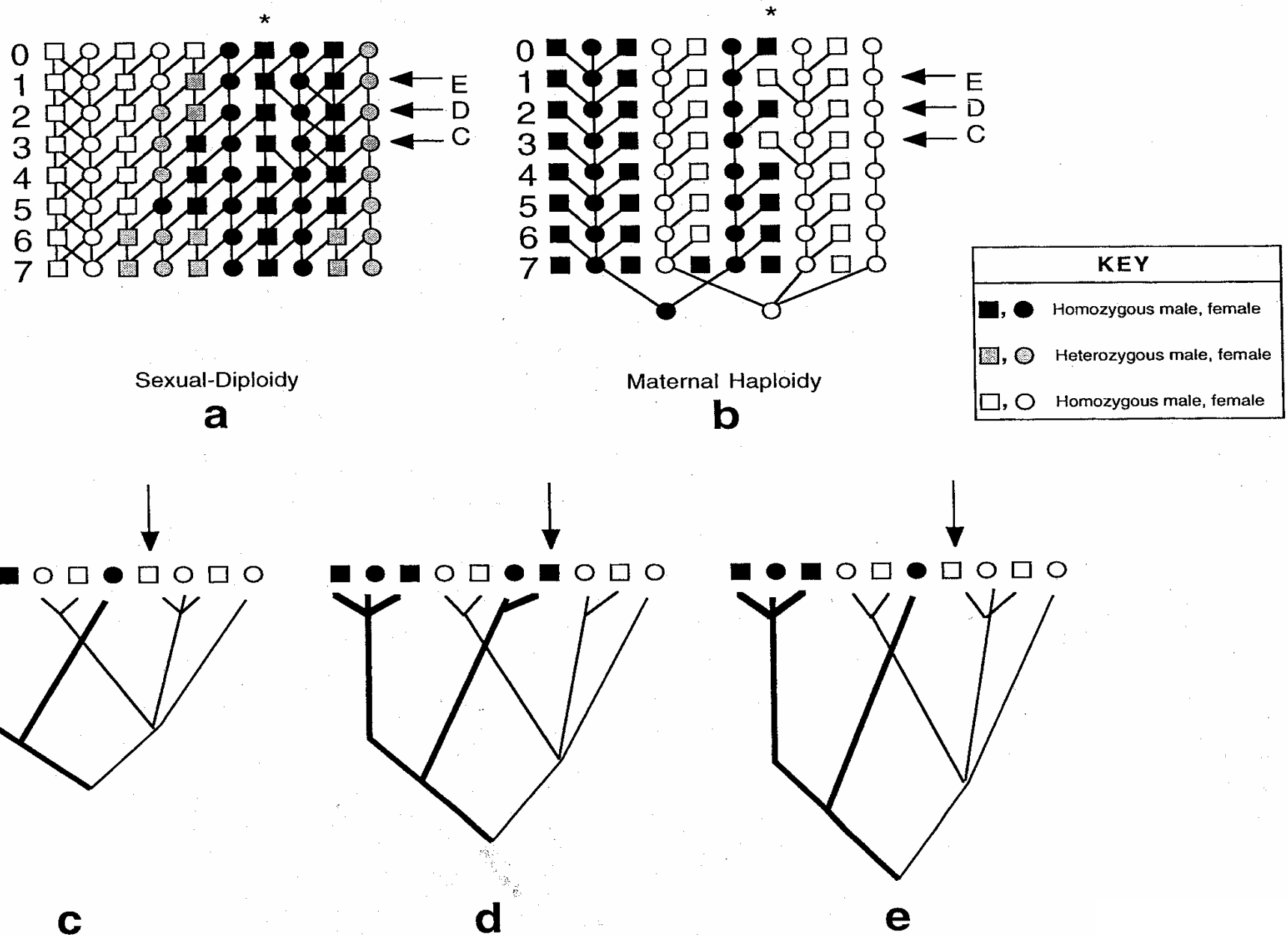


Reciprocal Monophyly

Caveats to tree-based methods

- Shoe-horning organisms into hierarchical relationships where none may exist: The limits of cladistic terms (monophyly, synapomorphy, etc.)
- Again, gene trees and species trees may not coincide

USGS
science for a changing world


*
0
1
2
3
4
5
6
7
E
D
C
Sexual-Diploidy
a
*
0
1
2
3
4
5
6
7
E
D
C
Maternal Haploidy
b
KEY
Homozygous male, female
Heterozygous male, female
Homozygous male, female
c
d
e

Caveats to tree-based methods

- Shoe-horning organisms into hierarchical relationships where none may exist: The limits of cladistic terms (monophyly, synapomorphy, etc.)
- Phylogeography TAKEN ALONE or performed uncritically may mislead

USGS
science for a changing world

Caveats to tree-based methods

- Shoe-horning organisms into hierarchical relationships where none may exist: The limits of cladistic terms (monophyly, synapomorphy, etc.)
- Phylogeography TAKEN ALONE or performed uncritically may mislead
- Concern over methods by which trees are generated (optimality criteria, nuclear vs. mitochondrial, phenetic versus cladistic combination of data sets)

USGS
science for a changing world

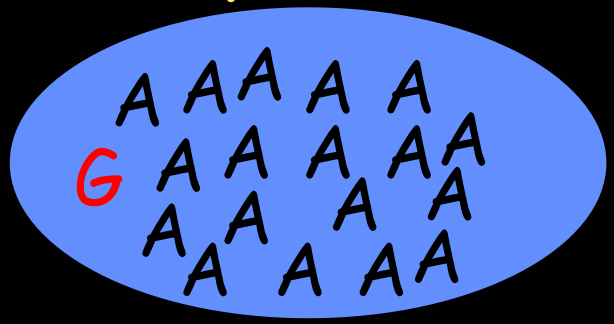
Caveats to tree-based methods

- Shoe-horning organisms into hierarchical relationships where none may exist: The limits of cladistic terms (monophyly, synapomorphy, etc.)
- Phylogeography TAKEN ALONE or performed uncritically may mislead
- Concern over methods by which trees are generated (optimality criteria, nuclear vs mitochondrial, combination of data sets)
- Ability to deal with non-molecular data
  - Relevance of genetic data to behavior
  - Evolutionary "fate" or "potential"

USGS
science for a changing world

Phylogenetic species and Population Aggregation Analysis (Davis and Nixon, 1992)

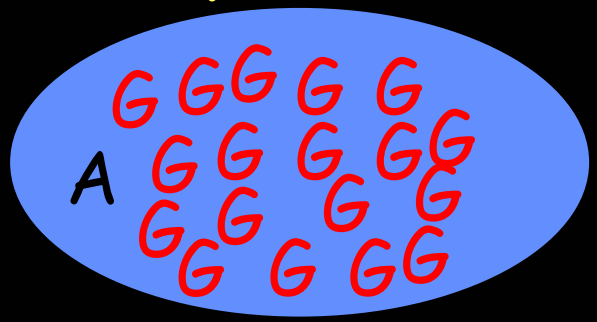
Population 1



 # Phylogenetic Species

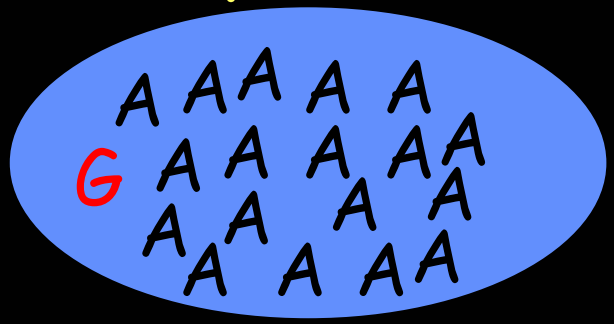
1

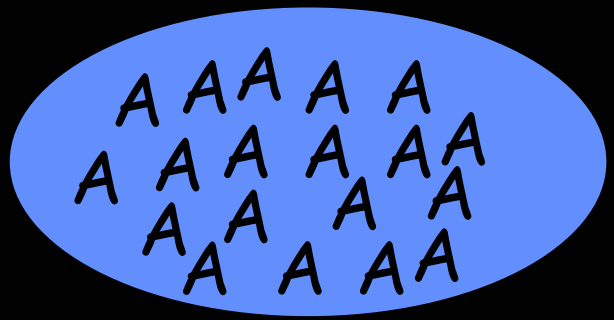
Population 1



Phylogenetic species and Population Aggregation Analysis (Davis and Nixon, 1992)

Population 1

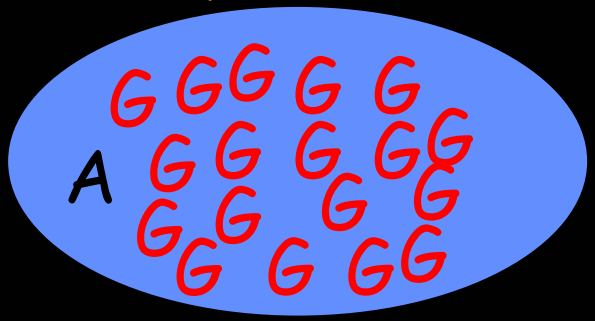


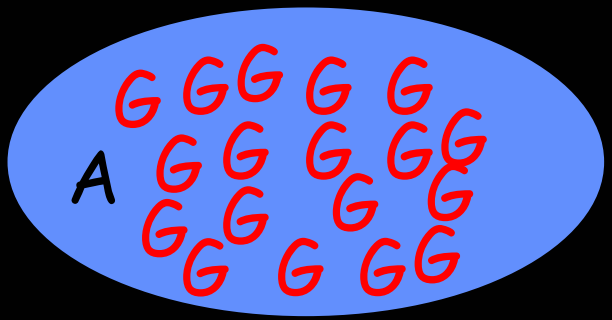


 # Phylogenetic Species

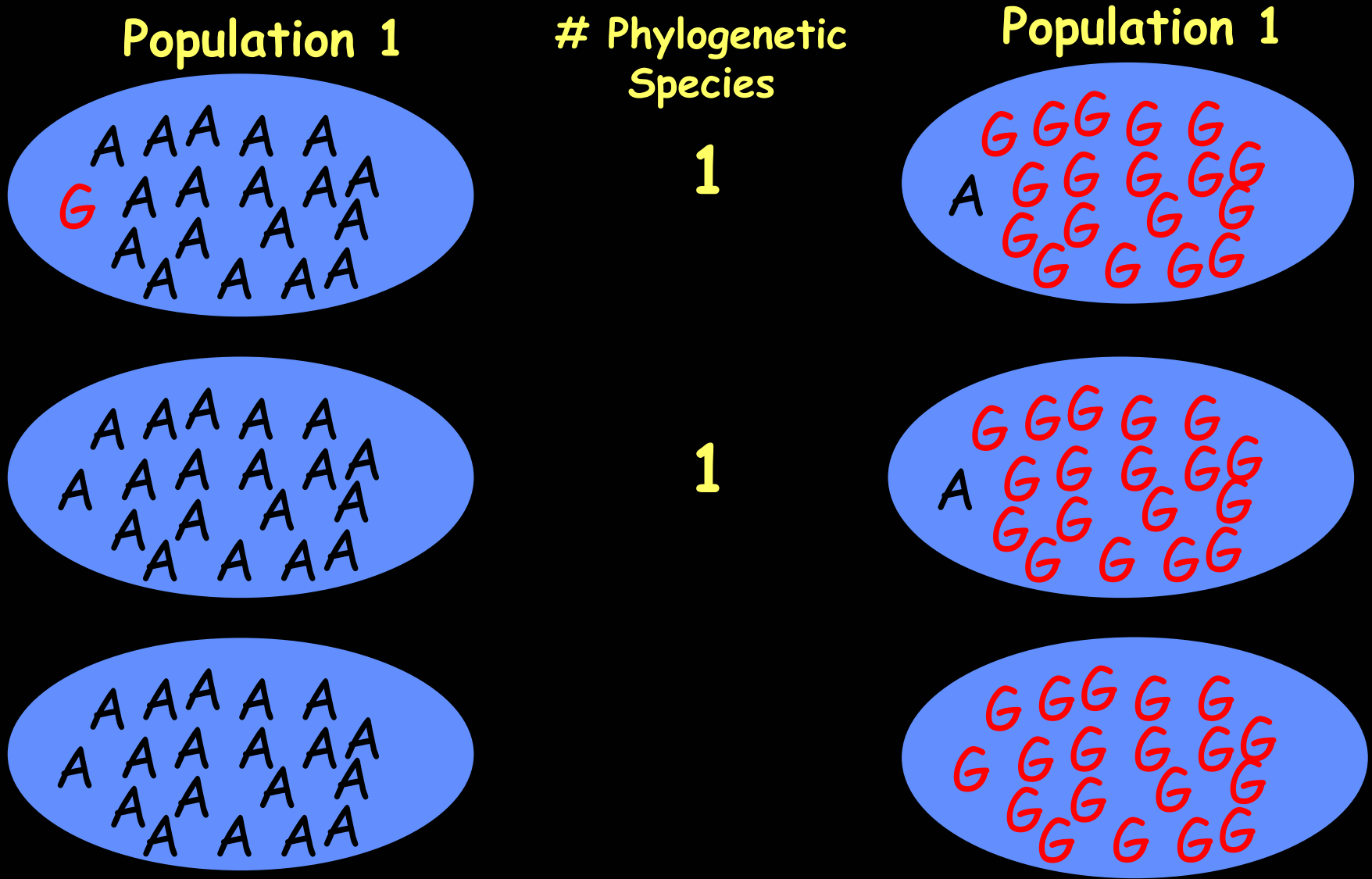
1

Population 1

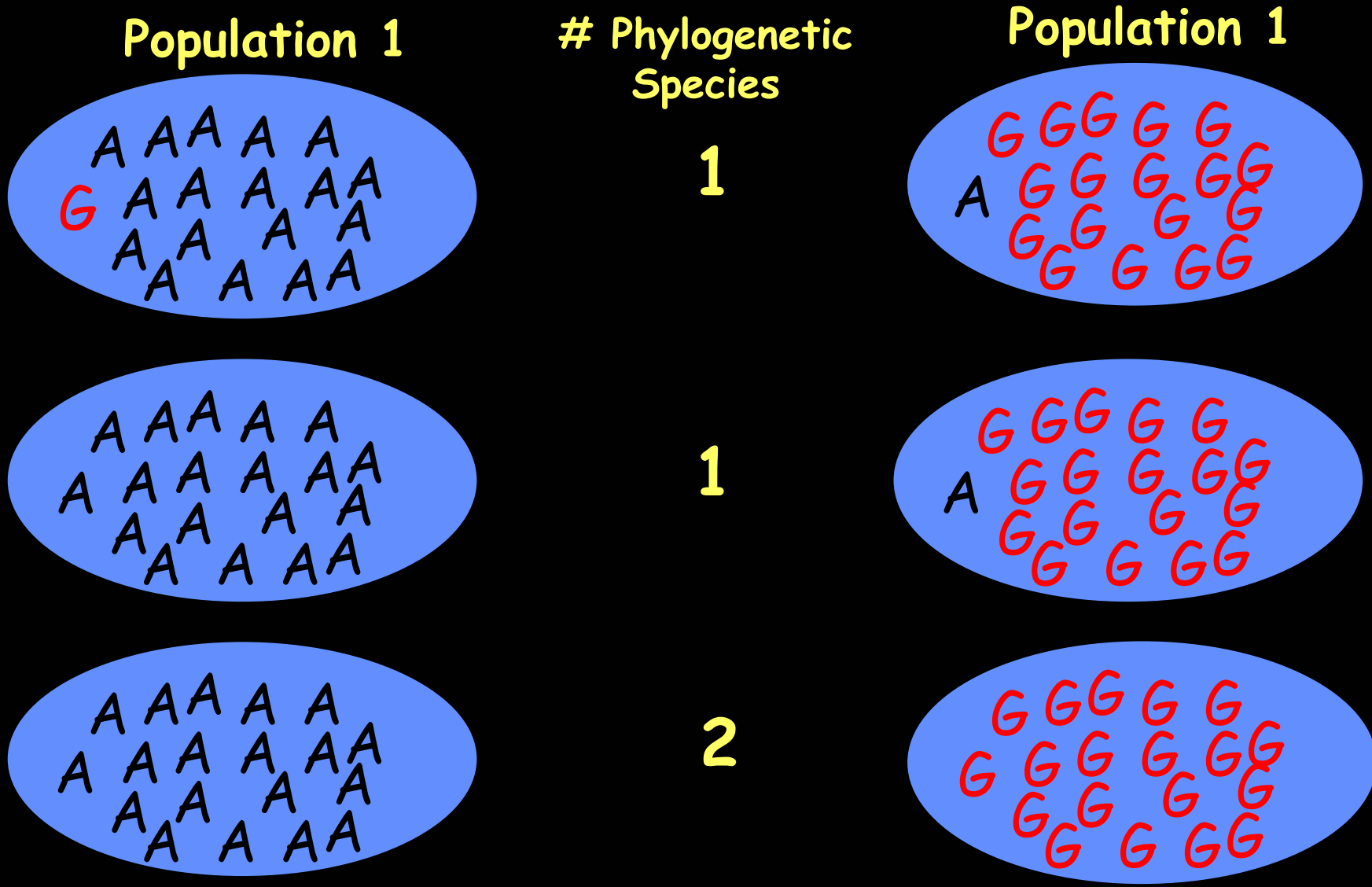




Phylogenetic species and Population Aggregation Analysis (Davis and Nixon, 1992)

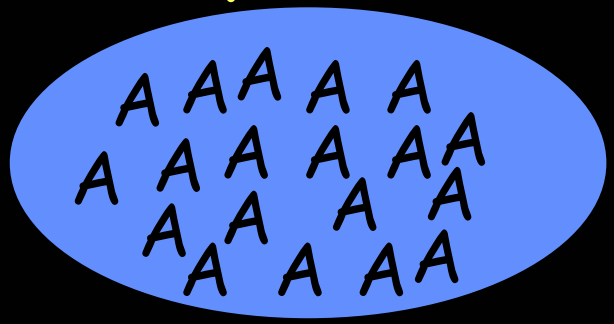


Phylogenetic species and Population Aggregation Analysis(Davis and Nixon, 1992)



Phylogenetic species and Population Aggregation Analysis (Davis and Nixon, 1992)

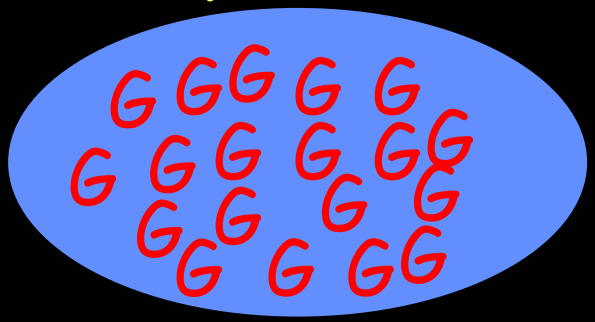
Population 1

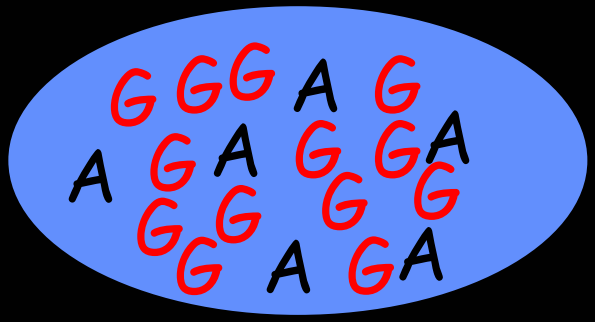


 # Phylogenetic Species

1

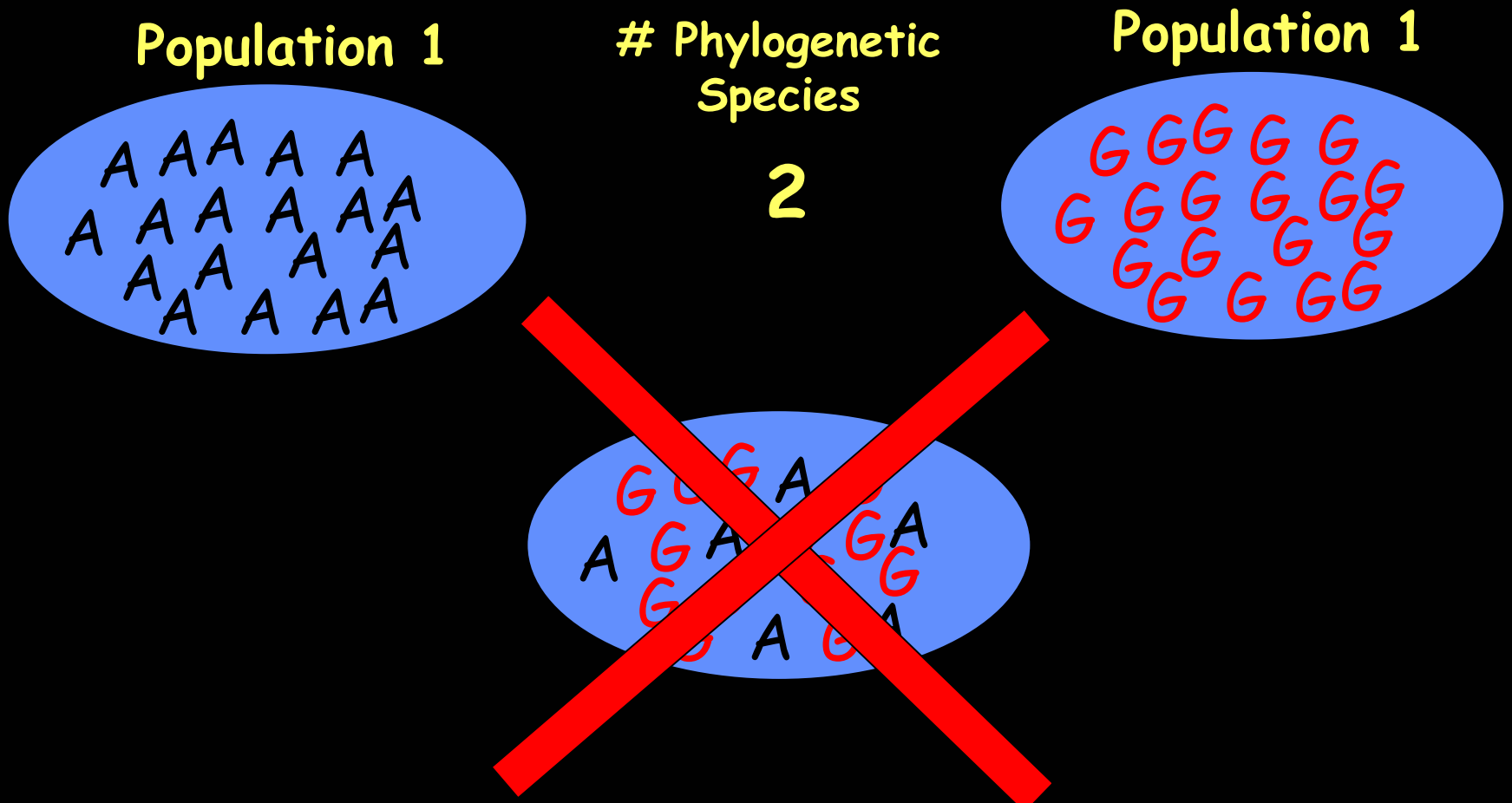
Population 1





USGS
science for a changing world

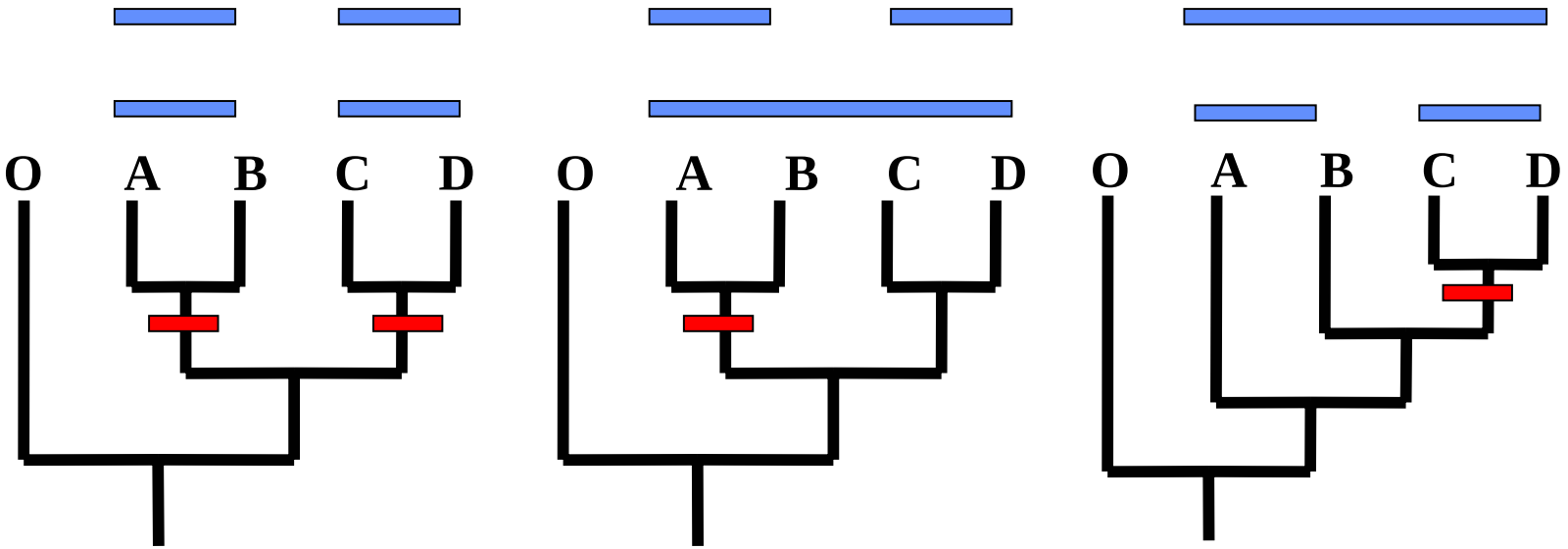
Phylogenetic species and Population Aggregation Analysis (Davis and Nixon, 1992)



USGS
science for a changing world

Tree-based

Character-based

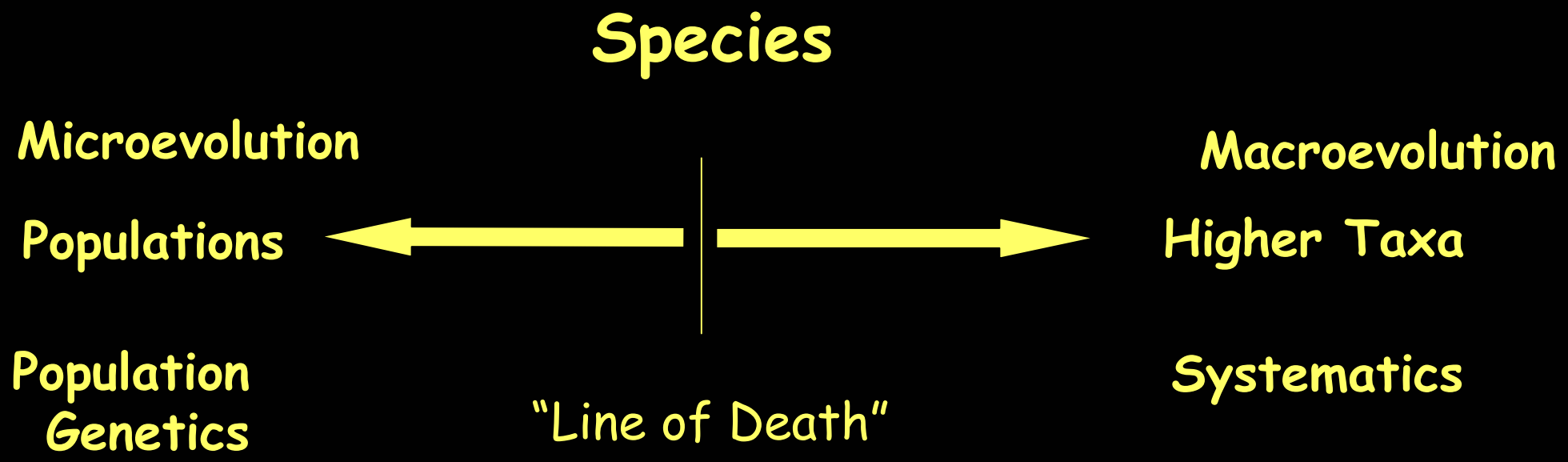


Reciprocal Monophyly + Character Support

Reciprocal Monophyly - Character Support

Non-reciprocal Monophyly

USGS
science for a changing world



After Davis and Nixon, 1992

USGS science for a changing world

Caveats to character-based methods

- Over-diagnosis or under-diagnosis as a result of poor sampling of individuals or genes

USGS
science for a changing world

Caveats to character-based methods

- Over-diagnosis or under-diagnosis as a result of poor sampling of individuals or genes
- Although these examples are highly simplified, multiple characters with congruent distributions (fixation patterns) are desirable. No one would name a new species on the basis of a single nucleotide polymorphism

USGS
science for a changing world

Caveats to character-based methods

- Over-diagnosis or under-diagnosis as a result of poor sampling of individuals or genes
- Although these examples are highly simplified, multiple characters with congruent distributions (fixation patterns) are desirable. No one would name a new species on the basis of a single nucleotide polymorphism
- Identification of the species boundary?

USGS
science for a changing world

Caveats to character-based methods

- Over-diagnosis or under-diagnosis as a result of poor sampling of individuals or genes
- Although these examples are highly simplified, multiple characters with congruent distributions (fixation patterns) are desirable. No one would name a new species on the basis of a single nucleotide polymorphism
- Identification of the species boundary?
- Proliferation of names?

USGS
science for a changing world

Implications of a character-based approach

- Solution to Woodger's dilemma (= "speciation" is instantaneous, corresponding to fixation)
- A single individual suffices to refute the hypothesis of character fixation
- Generalizable (all organisms have characters)
- Focus is on utility, i.e. on reconstructing the past, not future stability of lineages or species names
- Allows for the study of reproductive isolation without hanging up taxonomy & systematics

USGS
science for a changing world

THE ROYAL SOCIETY

FirstCite® e-publishing

Received 4 August 2003
Accepted 17 October 2003
Published online

The role of subspecies in obscuring avian biological diversity and misleading conservation policy

Robert M. Zink

Bell Museum, 100 Ecology Building, University of Minnesota, St Paul, MN 55108, USA (*rzink@biosci.cbs.umn.edu*)

Subspecies are often used in ways that require their evolutionary independence, for example as proxies for units of conservation. Mitochondrial DNA sequence data reveal that 97% of continentally distributed avian subspecies lack the population genetic structure indicative of a distinct evolutionary unit. Subspecies considered threatened or endangered, some of which have been targets of expensive restoration efforts, also generally lack genetic distinctiveness. Although sequence data show that species include 1.9 historically significant units on average, these units are not reflected by current subspecies nomenclature. Yet, it is these unnamed units and not named subspecies that should play a major role in guiding conservation efforts and in identifying biological diversity. Thus, a massive reorganization of classifications is required so that the lowest ranks, be they species or subspecies, reflect evolutionary diversity. Until such reorganization is accomplished, the subspecies rank will continue to hinder progress in taxonomy, evolutionary studies and especially conservation.

Keywords: mitochondrial DNA; subspecies; reciprocal monophyly; conservation

USGS science for a changing world

Subspecies: How do they fit in to a phylogenetic framework?

- They don't

USGS
science for a changing world